Macroeconomics

WARREN L. SMITH
Professor of Economics
The University of Michigan

1970

RICHARD D. IRWIN, INC. Homewood, Illinois 60430
IRWIN-DORSEY INTERNATIONAL, London, England WC2H 9NJ
IRWIN-DORSEY LIMITED, Georgetown, Ontario L7G 4B3

FIRST PRINTING, JULY, 1970
SECOND PRINTING, MAY, 1971
THIRD PRINTING, NOVEMBER, 1971
FOURTH PRINTING, JANUARY, 1974

Library of Congress Catalog Card No. 70–118183
PRINTED IN THE UNITED STATES OF AMERICA

To my wife
Ann Schwartz Smith

PREFACE

This book is intended to provide a fairly comprehensive coverage of modern macroeconomics. While considerable attention is given to individual macroeconomic relationships, the combination of these relationships to form a coherent explanation of the overall behavior of the economy is emphasized. Heavy use is made in many parts of the book of linear models, both static and dynamic. As the reader will see, the taste of the author runs heavily toward the use of algebra as an explanatory device. Many of the static models that are presented could have been formulated in more general terms, and the differential calculus could have been used to analyze them. In fact, however, very little would have been added by this approach, and the material would have been rendered very difficult for students whose knowledge of the calculus is shaky or nonexistent. The algebraic approach that is used is in some places quite elaborate, but any student with sufficient motivation and patience should be able to master it. The only place where the calculus is (almost unavoidably) used is in a few footnotes in Chapter 19, which deals with neoclassical growth theory. Even here, the student should be able to master the material without an extensive knowledge of mathematics.

The book is unabashedly Keynesian—or, more accurately, neo-Keynesian—in spirit. In Part II, which deals with aggregate demand and income determination and constitutes the heart of the book, a considerable amount of emphasis is placed on the way in which monetary forces interact with nonmonetary influences in determining the course of the economy. Chapter 14, which develops the working of the monetary sector in more detail than is customary in books of this kind, is an important element in rounding out the analysis.

Many persons have contributed to the development of this volume, either by assisting in the preparation of portions of the text, by reading all or portions of the manuscript and providing critical comments, or by using portions as text material in their classes and obtaining student reactions. Those who have been especially helpful in one or more of these ways include W. H. Locke Anderson, Robert S. Holbrook, Jay H. Levin, Harold M. Levinson, Johan A. Lybeck, William E. McFarland, Gail Pierson, Robert H. Rasche, and Ronald L. Teigen. I am also indebted to the secretarial staff of the Department of Economics, University of Michigan, for help in the preparation of the manuscript.

June, 1970 WARREN L. SMITH

CONTENTS

Chapter 1 INTRODUCTION

In this book, we shall be interested in the forces that affect the overall performance of the economy. That is, we shall study the determinants of such economywide measures of economic performance as the level and rate of growth of national income and product, the level of employment and unemployment, the general price level, and the nation's balance of international trade and payments.

The average citizen is often critically affected by the economy's performance. Perhaps the most dramatic illustration of this fact is the Great Depression of the 1930's, when for an entire decade the unemployment rate averaged nearly 20 percent—that is, one out of every five persons able and willing to work was unable to find employment—and the national output of goods and services was consistently far below the economy's productive potential. Severe economic distress was widespread, and many people became deeply disillusioned with our economic and political institutions. Since World War II the economy has performed much more effectively, and there has been no repetition of the experience of the 1930's. There have, however, been four brief recessions during the postwar era—in 1948–49, 1953–54, 1957–58, and 1960–61—and a prolonged period lasting from 1958 to 1963 during which the unemployment rate was persistently in excess of 5 percent.

The well-being of many people may also be significantly affected by a persistent rise in the price level. The United States has experienced several inflationary episodes since World War II. There was a sharp rise in prices when wage and price controls were removed at the end of the war. A second sharp inflationary outburst in 1950–51 was triggered by the outbreak of the Korean War in June 1950. There was also a period of more moderate but nevertheless worrisome inflation between 1955 and 1958. And recently we have again encountered a serious problem of persistent inflation. Starting from an unemployment rate of nearly 7 percent in early 1961, the economy began a prolonged period of expansion, which at the time this is being written (mid-1969) has continued without interruption for eight and a half years. For the first four and a half years or so—up to mid-1965—the expansion was smooth and orderly; the unemployment rate declined slowly but rather steadily, and prices remained quite stable. Beginning in late 1965, however, a rapid increase in military spending associated with the escalation of the conflict in Vietnam disrupted the orderly pattern of the expansion. Since that time unemployment has fallen to unusually low levels—less than 3.5 percent—but increasingly severe inflationary tendencies have developed.

The burdens imposed by inflation are much less obvious and more subtle than those caused by unemployment. Undoubtedly the burden is partly illusory: the housewife who sees the prices of things she buys rising from week to week believes that inflation is robbing her family of a portion of the real income earned by her husband; what she often fails to realize is that commonly the same forces that raise the prices she has to pay also raise her husband's money wages. Inflation does, however, impose real burdens on some people—those whose money incomes do not adjust in pace with rising prices and those who hold their wealth in forms whose money value remains constant. Moreover, inflation many cause distortions in the economy which reduce economic efficiency. And, finally, continuing inflation may make it difficult for a nation's goods to compete effectively in foreign markets, thereby leading to balance-of-payments difficulties.

Nearly everyone agrees that it would be highly desirable to have both low unemployment and a stable price level at the same time; however, the experience of recent years—not only in the United States but in many other countries as well—indicates that this combination is very difficult to achieve, because the same forces that reduce unemployment also tend to cause inflation.

Economic growth also has important implications for the average citizen. The living standard of the people of a country is best measured by income per head of the population. Clearly, the living standard measured in this way rises over time to the extent that the rate of growth of aggregate real income exceeds the rate of growth of the population. Over the long run the United States has experienced vigorous growth and a living standard that has risen until it is the highest in the world. There have, however, been periods when U.S. economic growth was sufficiently sluggish to be a source of serious public concern. This was especially the case in the 1930's when growth was negligible for an entire decade. But there was also a period of sluggish growth in the late 1950's—a period when the U.S. growth performance compared unfavorably with that of the Soviet Union, Japan, and several of the countries of Western Europe. During the 1960's U.S. growth has generally been vigorous once again, but interest in economic growth and its causes continues.

A final aspect of a nation's economy which is closely tied in with its overall economic performance is its balance-of-payments position. The forces that play on the balance of payments are very complex, but one relationship is fairly clear: when the economy is expanding vigorously and incomes are rising, imports of goods and services expand; if prices are also rising, the expansion of exports may be inhibited as the nation's competitive position in world markets deteriorates. Thus, by stimulating imports and inhibiting exports, vigorous economic expansion may push the balance of payments toward a deficit. Under our present international monetary system, a country which experiences a persistent deficit in its balance of payments will experience difficult problems as its reserves of gold and foreign exchange decline. Accordingly, there may be times—as has been the case for the United States during much of the 1960's—when a nation encounters difficulties in maintaining a healthy domestic economy and a viable balance-of-payments position at the same time.

It is to these important issues—employment and unemployment, the behavior of the price level, the rate of growth of income, the nation's balance of payments position, and how these aspects of a nation's overall economic performance are related to one another—that the analysis presented in this book is directed.

THE NATURE AND DEVELOPMENT OF MACROECONOMICS

The subject matter of this book—the study of the overall performance of the economy—is usually called "macroeconomics." This is to be contrasted with "microeconomics," which deals with the behavior of individual decision-making units such as households, business firms, and governmental units.[1]

Relation between Macroeconomics and Mircoeconomics

Since the actual decisions about production, consumption, employment, pricing, etc., are made by the microunits of the economy, the basic principles of economic theory are those which explain the behavior of these units. This being the case, it might appear that the best approach to the explanation of the behavior of the economy as a whole would be to explain the behavior of each decision-making unit and then combine these explanations into a theory of the entire economy. An approach along these lines, called "general equilibrium theory," has been developed at quite an abstract level. The study of general equilibrium theory is useful in developing an understanding of

[1] The distinction between macroeconomics and microeconomics is not as sharp as this statement suggests. Economists often study problems that extend beyond the individual decision-making unit but do not take in the entire economy. Examples of this kind would include a study of the coal industry, a study of the urban ghetto of a large city, or a study of the agricultural sector of the economy. Work of this kind is usually classified under the rubric of microeconomics. However, an analysis of a major geographic region—which might include only part of a country or, alternatively, more than one country—might employ many of the techniques of macroeconomics such as those developed in this book.

the relations between component parts of the economy and of the way in which the individual decision-making units fit together to form a coherent whole.

However, the general equilibrium approach is far too detailed and complex to be of practical use in explaining the behavior of the economy as a whole. A more practical approach, which is followed to a considerable extent in this book, is to divide the economy into sectors, such as the household sector, the business sector, and the government sector; develop explanations of the relevant behavior of these sectors; and combine these explanations to arrive at the behavior of the economy as a whole. The microeconomic theory that has been developed to explain the behavior of individual decision-making units cannot generally be applied directly to sectors—for example, the theory of the individual household cannot be directly applied to the household sector. The reason for this is that there is often an "aggregation problem"—the result of each unit in a sector behaving individually in accordance with a certain theory will not be the same as the result that would be obtained if the entire sector acting as a unified whole behaved in accordance with the theory. (A very simple analogy: the sum of the square roots of a series of numbers is not equal to the square root of the sum.) Nevertheless, microeconomic theory can often be a fruitful source of hypotheses that can be used with suitable modifications to explain sectoral behavior. Indeed, the student will encounter a number of illustrations of this procedure in this book.

The Development of Macroeconomics

Prior to 1930, most economists believed that a free enterprise economy contained powerful built-in tendencies toward full employment. It was recognized that departures might occasionally occur, but the generally accepted view was that these departures would seldom be great and that when they did occur, automatic forces could be depended upon to return the system rapidly to full employment. Thus, the attitude toward the problem of unemployment was a combination of complacency and fatalism: complacency, because it was regarded as unlikely that unemployment would be a serious problem; fatalism, because when unemployment did occur the prevalent view was that there was little to be done about it other than to let natural forces work without interference. Monetary policy conducted by the central bank was viewed as the only legitimate form of macroeconomic policy. Even here some economists thought—quite erroneously—that if proper rules were established governing the lending behavior of the banks "money would manage itself" without discretionary action by the central bank or the government. But as the 1920's progressed, more and more economists became convinced of the desirability of deliberate monetary management by the central bank. However, it was generally felt that the appropriate objectives of central bank policy were the maintenance of a stable price level or the achievement of balance-of-payments equilibrium.

These traditional and complacent views were completely shattered by the Great Depression of the 1930's. Massive unemployment made its appearance—not only in the United States but in most other countries as well. Moreover, years went by without either the automatic forces stressed by economists or the rather fumbling policy actions by governments achieving much success in reducing unemployment.

Clearly a reformulation of macroeconomic theory was badly needed. Such a reformulation was provided by John Maynard Keynes in *The General Theory of Employment, Interest, and Money* published in 1936. The student should understand that Keynes was a strong supporter of capitalism. Indeed, in his view the single major defect of capitalism in modern industrialized society was that—contrary to the view of most economists of his time—the system did not contain an effective built-in mechanism for keeping it close to full employment or returning it promptly to full employment if a departure should occur. In the absence of such a mechanism, Keynes contended that it was necessary for the government to take responsibility through its economic policies for the maintenance of full employment. Apart from this important amendment, Keynes accepted the established precepts of neoclassical economics.

Another revolutionary aspect of Keynes' *General Theory* was the demonstration it contained that the government could affect the overall level of economic activity through adjustments in its fiscal position—that is, by changing the level of its own expenditures or by changing its tax laws to alter its revenue collections. Before Keynes, some pamphleteers and amateur economists had suggested direct government spending as a cure for unemployment, but such ideas were generally looked upon with disdain by professional economists. To the extent that economists had come to recognize the need for the government to take action to combat unemployment prior to the *General Theory,* they generally felt that central bank monetary policy was the only instrument that could be used for this purpose. Keynes, on the other hand, took the position that under the conditions of severe unemployment that existed at the time his book was written, monetary policy was likely to be almost totally ineffective, leaving fiscal policy as the only effective instrument for fighting unemployment.

Keynes' *General Theory* was certainly not the last word in macroeconomics. The book was to a considerable extent a tract of the times, designed to explain the chronic unemployment of the 1930's and to provide the basis for formulating effective policies to combat that unemployment. Consequently, some alterations and additions were needed to adapt it to the analysis of a normally prosperous and secularly growing economy. Nevertheless, Keynes' work provided a solid foundation for the development of modern macroeconomics. Accordingly, the analysis contained in this book is decidedly Keynesian in spirit—or perhaps it might be better to call it "neo-Keynesian," since an attempt is made to incorporate many of the advances that have been made since the *General Theory* was written.

Views regarding the role of government in relation to the economy have changed greatly since the early 1930's, partly as a result of the work of Keynes and his successors. It is now generally agreed—both here and abroad—that the fiscal and monetary policies followed by the government can have important effects on the economy and that the government has a responsibility to follow policies designed to achieve a high standard of overall economic performance. In the United States this responsibility was specifically written into law in the Employment Act of 1946, which reads in part:

> The Congress hereby declares that it is the continuing responsibility of the Federal Government to use all practicable means . . . for the purpose of creating and maintaining . . . conditions under which there will be afforded useful employment opportunities . . . for those able, willing, and seeking to work, and to promote maximum employment, production, and purchasing power.

Despite a general acceptance of the view that the government has a responsibility for the performance of the economy, there is still considerable controversy about various aspects of macroeconomics. In part, this controversy involves value judgments rather than issues that are amenable to economic analysis. Sometimes two or more objectives, each of which is desirable, may come into conflict. For example, measures that reduce unemployment may generate inflationary pressures. In such circumstances, a question arises: Which is better for the nation, reducing unemployment or keeping the price level stable? The economist may be able to say something useful about the consequences of stressing one of the conflicting objectives as compared with the other. But even if there is agreement on the consequences, there can still be legitimate differences of opinion based on social and political philosophy—differences which cannot be ultimately resolved by economic reasoning alone.

Economists also disagree about issues of economic analysis. While great progress has been made toward improved understanding of macroeconomics in recent years as a result of extensive theoretical work and empirical investigation, substantial uncertainties and gaps in our knowledge remain. For example, while most economists would agree that both fiscal and monetary policy can have significant effects on economic activity, there is a substantial difference of opinion among competent investigators regarding their relative importance. The sources of some of the disagreements of this kind will come out in the course of this book.

THE USES OF ECONOMIC MODELS

Since much of the analysis presented in this book—and indeed nearly all theorizing and analysis in economics—involves the use of so-called "models" of economic behavior, some preliminary discussion of model building will be useful.

The Model-Building Approach

Models are merely simplifications or idealizations of reality. The reason for their extensive use in economics is simply that economic reality is so complex and involves such an incredible welter of interrelated variables that a direct approach to it is almost always impracticable. Take, for example, the forces determining the general level of prices or the purchasing power of money. This is clearly a meaningful concept and in fact one about which there is a great deal of legitimate public concern. However, the general price level is difficult to measure—in fact, there is no wholly satisfactory measure available—because individual prices move differently and the number of ways in which they can be combined to get a measure of the average movement is literally endless.[2] Moreover, once the problem of measurement has been resolved by selecting some agreed-upon average, the number and complexity of the forces having some effect on the level of prices so measured is colossal. Some of these forces are very important while others are of trivial significance. Any effort to develop an attack on the problem that will take all of the forces into account is clearly out of the question. The only feasible approach is to construct a model of the economic process which abstracts from the innumerable minor forces and focuses on a few major ones.

Models are sometimes developed to deal with rather specific questions. Our approach will be different in this book. In the next few chapters we shall develop a model which is rather general in its applicability to the kinds of problems we are concerned with—although it will, of course, have its limitations, which we shall explain in due course. Our purpose is to provide the student with a reasonably general framework of analysis to guide and discipline his thinking concerning the matters we are interested in, one which is adaptable within limits to dealing with a variety of problems and policy issues. Our method will be to start with rather simple relationships and then build in successive complications gradually.

---

[2] The problem of measuring changes in the price level is taken up at some length in Chapter 4.

THE USES OF ECONOMIC MODELS

Since much of the analysis presented in this book—and indeed nearly all theorizing and analysis in economics—involves the use of so-called "models" of economic behavior, some preliminary discussion of model building will be useful.

The Model-Building Approach

Models are merely simplifications or idealizations of reality. The reason for their extensive use in economics is simply that economic reality is so complex and involves such an incredible welter of interrelated variables that a direct approach to it is almost always impracticable. Take, for example, the forces determining the general level of prices or the purchasing power of money. This is clearly a meaningful concept and in fact one about which there is a great deal of legitimate public concern. However, the general price level is difficult to measure—in fact, there is no wholly satisfactory measure available—because individual prices move differently and the number of ways in which they can be combined to get a measure of the average movement is literally endless.[2] Moreover, once the problem of measurement has been resolved by selecting some agreed-upon average, the number and complexity of the forces having some effect on the level of prices so measured is colossal. Some of these forces are very important while others are of trivial significance. Any effort to develop an attack on the problem that will take all of the forces into account is clearly out of the question. The only feasible approach is to construct a model of the economic process which abstracts from the innumerable minor forces and focuses on a few major ones.

Models are sometimes developed to deal with rather specific questions. Our approach will be different in this book. In the next few chapters we shall develop a model which is rather general in its applicability to the kinds of problems we are concerned with—although it will, of course, have its limitations, which we shall explain in due course. Our purpose is to provide the student with a reasonably general framework of analysis to guide and discipline his thinking concerning the matters we are interested in, one which is adaptable within limits to dealing with a variety of problems and policy issues. Our method will be to start with rather simple relationships and then build in successive complications gradually.

---

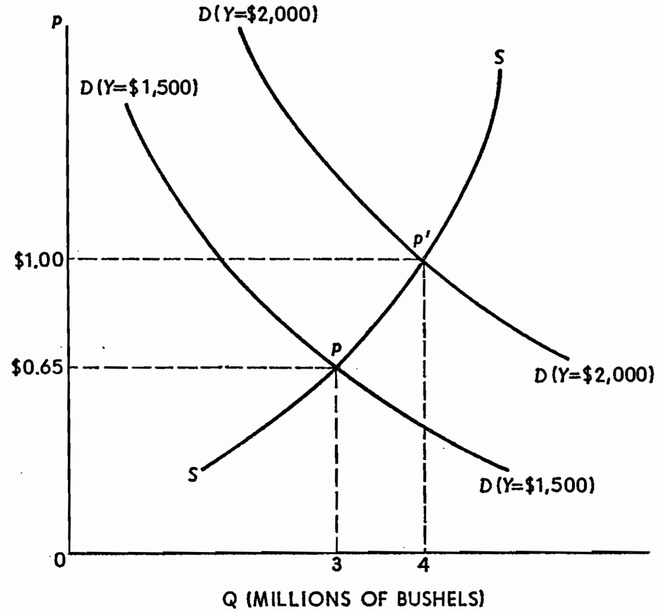
[2] The problem of measuring changes in the price level is taken up at some length in Chapter 4.

three variables Q_d, Q_s, and p. These three variables may be called the *endogenous* variables of the model, meaning that they are the variables that are determined within the model. The variable Y, on the other hand, is an *exogenous* variable in this model; that is, the model does not contain sufficient relationships to determine it, and its value must be given to us before we can use the model to determine the other three variables.

A possible solution of this model is depicted in Figure 1–1. Here the

FIGURE 1–1

Hypothetical Demand and Supply Relationships for Potatoes



curve SS is the suppy function (S) showing how the quantity supplied (measured horizontally) varies as the price (measured vertically) varies. The curve marked $D(Y = \$1{,}500)$ is a demand schedule showing how quantity demanded varies as price varies when income is held constant at a level of $1,500. It is apparent in this case that if income per capita is $1,500 per year, the price of potatoes will be $0.65 per bushel and the amount demanded (and supplied) will be three million bushels per year.

This analysis is *static*—that is, given the nature of the functional relationships involved and given the value of the exogenous variable, Y, the model determines the equilibrium values of price and quantity. As long as the underlying conditions and the value of income remain unchanged, the price will continue to be $0.65 and the rate of production and consumption will continue to be three million bushels.

We may extend the use of this model a bit further by considering what happens if the exogenous variable, income, changes or if a shift in either the demand curve or the supply curve occurs due to a change in some other underlying factor such as the tastes of consumers. For example, suppose that for some reason the level of income rises to \$2,000. When this happens, the amount of potatoes that will be demanded at each price increases so that the demand curve in Figure 1–1 shifts. Let us suppose it takes up the position indicated by the curve designated $D(Y = \$2{,}000)$. When this happens the new equilibrium point is four million bushels at a price of \$1.00. This kind of analysis, which shows what happens to the *equilibrium* values of the endogenous variables when there is a specified change in an exogenous variable or in a functional relationship, is known as *comparative statics.* Since much of economic theory is concerned with the effects produced by changes in underlying conditions, comparative statics is an important kind of analysis.

In drawing Figure 1–1, it was assumed that we somehow knew the exact nature of the functional relationships of supply and demand and were therefore in a position to draw up precise supply and demand curves. This would frequently not be the case, although under some circumstances econometric investigations might permit us to estimate the functional relationships with tolerable confidence.[3] However, even if we do not know exactly what the functional relationships are, we may still be able to say something useful concerning the effects of *changes* in exogenous variables or underlying conditions. For example, it is apparent from Figure 1–1 that a shift to the right of the demand curve will cause both the equilibrium price and the equilibrium quantity to increase provided the supply curve has a positive slope (i.e., slopes upward to the right). Thus, if we knew on the basis of a priori reasoning, supported perhaps by some empirical knowledge or familiarity with the market, that (*a*) an increase in income would cause a rise in demand for potatoes and (*b*) that the supply curve of potatoes had a positive slope, we might be able to predict with some confidence that a rise in income would cause a rise in the price of potatoes and in the quantity produced and consumed. Moreover, it is apparent that, other things being the same, the rise in price will be greater and the increase in quantity smaller the steeper the slope of the supply curve. Thus, we may be able to forecast the *direction* of effects and estimate their general *order of magnitude* without knowing the exact nature of the functional relationships involved.

It is important for the student to recognize that the supply and demand analysis of the type we have been discussing is a *model* rather than a universally applicable principle. A model of this kind can be derived formally from the behavior of the underlying decision-making units on the

---

[3] Econometric models are discussed briefly later in this chapter. See pp. 17–20.

assumption that a high degree of competition prevails. Thus, the more competitive the market being studied the more useful is this model. Strictly speaking it cannot be used at all in dealing with markets characterized by monopoly, because in such markets there is not a simple relationship between the price and the amount that will be supplied, so that a supply curve of the type employed in this model cannot be constructed.

One final point may be noted. When we are studying a small part of the economy such as the potato market, it is possible to treat total income as an exogenous variable. However, if we were studying a major sector of the economy—such as agriculture as a whole—this would not be permissible, since a change of significant magnitude in the agricultural sector would ordinarily cause a change in income which would be too large to be overlooked. Thus, in analysis of a major sector of the economy, it is necessary to treat income as an endogenous variable and to introduce sufficient additional relationships to determine it within the model. Even in the case of the potato market, a change in the price or production of potatoes will have an effect on the level of national income; however, for so small a sector of the economy such indirect effects will be rather unimportant, and it may be permissible to overlook them. In general, what is an exogenous variable for one model may be an endogenous variable in another model designed to deal with a different kind of problem.

Dynamic Models and Stability of Equilibrium

Comparative statics is a useful technique of economic analysis, but it does have its limitations. Although it permits us to make predictions about the changes which will occur in the equilibrium values of the endogenous variables as a result of specified changes in the exogenous variables or in the structure—i.e., the functional relationships—of the model, it tells us nothing about the *process* or *path* by which the variables move from one equilibrium position to the other. In fact, it does not even permit us to be sure that the new equilibrium position will be reached at all. *Dynamic analysis,* in cases where it can be applied, does permit us to study the time path of movement of the variables in some detail.

To illustrate dynamic analysis, let us revert to our example of the potato market. We will begin with a static model, but this time make the model specific by assuming linear relationships and giving them explicit numerical values.[4] Our model is as follows:

$$
\begin{aligned}
Q_d &= 40 - 3p + 0.02Y \\
Q_s &= 40 + 2p \\
Q_d &= Q_s
\end{aligned}
$$

[4] A linear relationship is one in which all the variables enter with constant coefficients or multipliers, such as $2x + 3y = 8$ or $0.14x - 3.12y = 48.3$.

If we assume, to start with, that Y is \$1,500, the first equation becomes

$$Q_d = 70 - 3p$$

since 0.02 times \$1,500 is 30, and 30 added to 40 gives us 70. Since by the last equation $Q_d = Q_s$, we can denote both by Q, and our model becomes

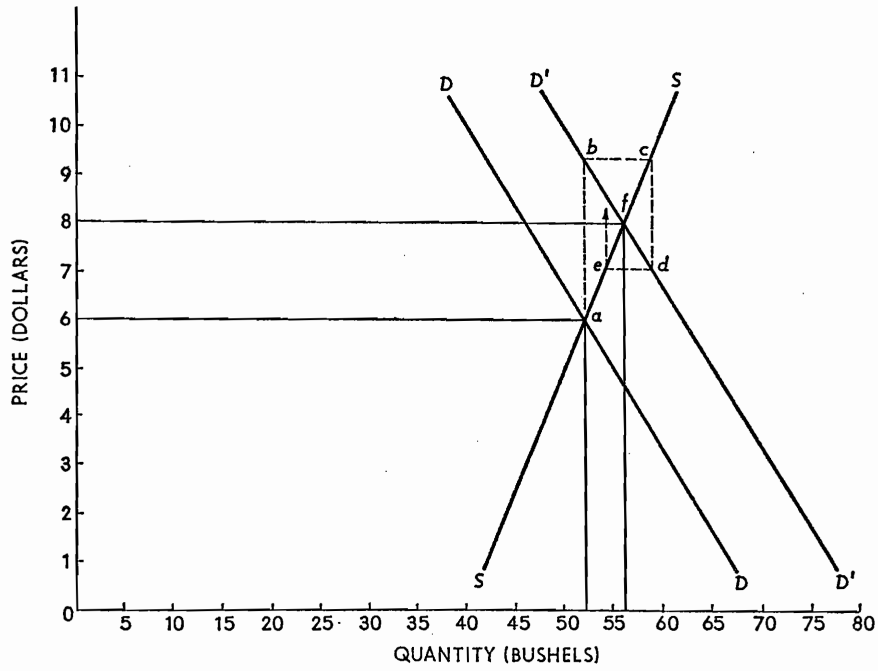
$$Q = 70 - 3p$$
$$Q = 40 + 2p$$

This can easily be solved, giving the values $Q = 52$ and $p = 6.0$.[5]

In Figure 1–2 the equation $Q = 70 - 3p$ is depicted by the demand

FIGURE 1–2

Hypothetical Dynamic Model of the Potato Market



curve DD, and the equation $Q = 40 + 2p$ is depicted by the supply curve SS. These two lines intersect at point a, where $p = 6$ and $Q = 52$, the same solution (of course) as that we obtained by algebra. In other words, when income is \$1,500, the equilibrium values of price and quantity are \$6 and 52 bushels, respectively.

[5] We can substitute $40 + 2p$ for Q in the first equation, which gives us $40 + 2p = 70 - 3p$, or $5p = 30$, or $p = 6$. Substituting $p = 6$ in either equation and solving for Q gives us $Q = 52$.

Now suppose income rises to \$2,000. This gives us a new demand schedule which is[6]

$$Q_d = 80 - 3p$$

and the model becomes

$$Q = 80 - 3p$$
$$Q = 40 + 2p$$

Solving this in the same way as before, we obtain $Q = 56$, $p = 8$. The new demand schedule $Q_d = 80 - 3p$ is depicted in Figure 1–2 by the curve $D'D'$, and this schedule intersects the supply schedule (which is unchanged) at point f, where $p = 8$ and $Q = 56$. Thus, the new equilibrium price and quantity, corresponding to income of \$2,000, are \$8 and 56 bushels, respectively.

Thus far, the model has been entirely static. Let us now introduce time into the picture in such a way as to specify how buyers and sellers adjust their actions to changing conditions. We shall suppose that the quantity demanded in a particular period depends upon the price and income in that period, or

$$Q_d(t) = 40 - 3p(t) + 0.02Y(t)$$

which should be read as, "Quantity demanded in period t is equal to 40 minus 3 times price in period t plus 0.02 times income in period t." On the supply side, on the other hand, we shall introduce a time lag, so that the quantity supplied in a particular period depends upon the price realized in the previous period. That is,

$$Q_s(t) = 40 + 2p(t - 1)$$

which should be read as "Quantity supplied in period t equals 40 plus 2 times price in period $t - 1$."

Thus, our model becomes

$$Q_d(t) = 40 - 3p(t) + 0.02Y(t)$$
$$Q_s(t) = 40 + 2p(t - 1)$$
$$Q_d(t) = Q_s(t)$$

This model depicts in a very simplified way a phenomenon that is sometimes found (although complicated and modified by other factors) in agricultural markets.[7] Producers of potatoes are assumed to base their decisions about the amount of their production on the price that prevails at the time the crop is

[6] Since 0.02 times 2,000 is 40, and 40 added to 40 gives us 80.

[7] Possible complicating factors include nonlinearities in the supply and demand schedules, variable and uncertain weather conditions, the possibility of storing the commodity from one crop season to another, the fact that not all plantings occur at the same time, the fact that suppliers may learn from experience, and so on.

planted. The crop is harvested some months later and thrown on the market for what it will bring. Thus, a new price is determined, and a new crop is planted based on this price.

The first thing to be noted is that the *equilibrium* solution for this model is the same as that for the simpler static model. To see this, let us suppose, as in our earlier analysis, that income is $1,500 and remains at that level from period to period. Since by the third equation, $Q_d(t) = Q_s(t)$, we can denote both by $Q(t)$ and drop the third equation. Doing this and substituting $1,500 for $Y(t)$, we obtain

$$Q(t) = 70 - 3p(t)$$
$$Q(t) = 40 + 2p(t - 1)$$

The equilibrium solution calculated above for the static equivalent of this model was $Q = 52$ and $p = 6$. Let us suppose that these values somehow come to prevail in a particular period. Then, by the second equation above, the quantity supplied in the next period will be 52 (40 plus 2 times 6), and when the 52 is thrown on the market, by the first equation, the price it will bring will be $6[$52 = 70 - 3p(t)$, $3p(t) = 18$, $p(t) = 6$]. Thus, once the values $Q = 52$ and $p = 6$ are achieved they will be *self-perpetuating* in the absence of some exogenous change in, say, income. That is to say, $Q = 52$ and $p = 6$ are equilibrium values. The equilibrium is satisfactorily depicted as in the static case by the intersection of the *SS* and *DD* curves in Figure 1–2.

Suppose now that income rises to $2,000 and remains at that level indefinitely. The dynamic model corresponding to this case can be depicted, according to the same reasoning used earlier, by

$$Q(t) = 80 - 3p(t)$$
$$Q(t) = 40 + 2p(t - 1)$$

By the method used above, the student can demonstrate for himself that the new equilibrium values of Q and p are 56 units and $8, respectively. Thus, as in the static case, the new equilibrium is depicted by the intersection (point *f*) of the *D′D′* and *SS* curves. However, the *process* by which the market arrives at the new equilibrium, starting from the old one, is rather complicated.

In the first period after the demand curve rises from *DD* to *D′D′*, suppliers continue to base their production on the old price of $6 and thus continue to produce 52 bushels. However, when this is thrown on the market under the new conditions, the price rises to $9⅓ or $9.33 [$52 = 80 - 3p(t)$, or $3p(t) = 28$, or $p(t) = 9.33$], which is above the new equilibrium price of $8. This leads producers to base their production plans on a price of $9.33, so that they produce 58.67 bushels [$Q(t) = 40 + 2(9.333)$], which is above the new equilibrium quantity of

TABLE 1–1

Time Path of Price and Quantity Following Change in Income in Cobweb Model

Period	Point in Figure II	Price	Quantity
0	a	$6.00	52.00
1	b	9.33	52.00
2	c	7.11	58.67
3	d	8.59	54.22
4	e	7.60	57.19
.	.	.	.
.	.	.	.
.	.	.	.
.	.	.	.
n	f	8.00	56.00

/D/ – demand function, $Q(t) = 80 - 3p(t)$.

/S/ – supply function, $Q(t) = 40 + 2p(t-1)$.

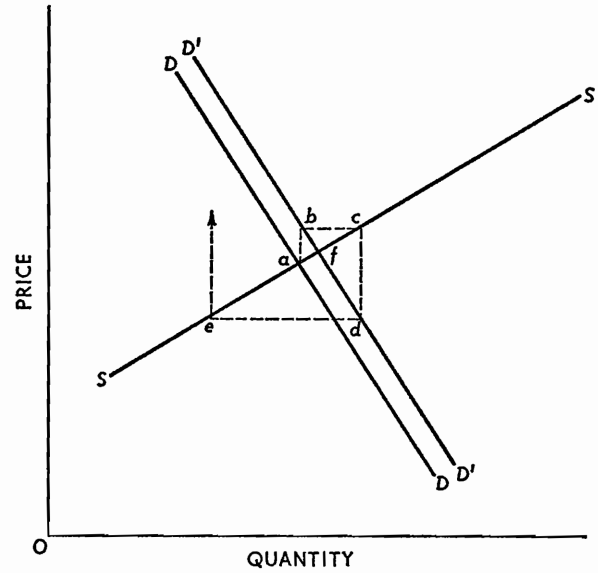
56. When 58.67 bushels are thrown on the market in the next period, the price is driven down to $7.11 (below the equilibrium price), and this leads suppliers to produce 454.22 bushels (below the equilibrium quantity). The results are summarized in Table 1–1, in which the arrows indicate the direction of causation and the enclosed letters /*D*/ and /*S*/ indicate whether the causation is due to the demand relation or the supply relation, respectively.

In Figure 1–2, price and quantity move along the spiral path *abcde.* It is apparent that the spiral is contracting and that the market is moving toward the new equilibrium at *f*. This can also be seen from Table 1–1, which shows that price and quantity are overshooting the mark first in one direction and then in the other in relation to the new equilibrium, but that the extent of the "overshoot" is continually growing smaller. Price and quantity are oscillating around the new equilibrium but the oscillations are "damped," that is, they are continually growing smaller and the market is gradually approaching the new equilibrium.

In a case of this kind, the system is said to be *stable.* That is, when the market is thrown out of equilibrium—as by the shift in demand in the above case—the forces at work are such as to push it toward the new equilibrium.

FIGURE 1–3

Illustration of an Unstable Dynamic Market Model



The stability of this type of market depends upon the slopes of the demand and supply curves relative to one another. When, as in the above example, the slope of the supply curve is greater (i.e., steeper) than the slope of the demand curve, the market is stable.[8] On the other hand, if, as in Figure 1–3,

[8] The comparison is between the *absolute values* of the slopes without regard to sign. Since price is measured on the vertical axis and quantity on the horizontal axis (as is conventional in economics), the slope refers to the change in price associated with a given change in quantity. Thus, the equation of the static demand curve *DD* in Figure 1–2 ($Q = 70 - 3p$) can be solved for price as a function of quantity, yielding $p = 70/3 - 1/3Q$. In this form, the equation indicates that an increase of one bushel in the quantity demanded is associated with a reduction of one third of a dollar (or 33 cents) in price; i.e., the slope of the demand curve is $-1/3$. Similarly, the supply equation ($Q = 40 + 2p$) can be solved for p, yielding $p = -20 + 1/2Q$, which indicates that an increase of one bushel in the quantity supplied is associated with an increase of one half of a dollar (or

the slope of the demand curve is greater than that of the supply curve, the market will be unstable. In Figure 1–3, the market was originally in equilibrium at point *a.* As in our earlier example, a shift of the demand curve from *DD* to *D′D′* throws the market out of equilibrium. The new equilibrium of demand and supply is at point *f*. As before, the market moves along a spiral path *abcde* and so on, but this time the spiral continuously expands instead of contracting and the market moves further and further from the new equilibrium. The market is thus *unstable* in the sense that instead of moving toward a new equilibrium it moves further and further away from it.[9]

Dynamic analysis has considerable advantages over static analysis. It enables us to analyze economic movements in some detail, and, in particular, it permits us to consider how the economy operates when it is out of equilibrium. Unfortunately, however, the technical complications are such that the range of problems to which we can apply it is somewhat limited. We shall make occasional use of dynamic analysis in the ensuing chapters of this book, but many of the problems with which we must deal are so complex that a dynamic treatment is impracticable. Consequently, much of our analysis will be of the static—or comparative static—variety, with such dynamic overtones as we can introduce in a nontechnical manner.

Theoretical versus Econometric Models

Econometrics is an approach to economic problems which, through a combination of economic theory and statistical analysis, attempts to quantify and measure economic relationships. Figure 1–4 illustrates econometric analysis by means of a very simple hypothetical example. Suppose an investigator sets out to find the demand curve for, say, potatoes in a particular market. He collects data on the price of potatoes and the quantity consumed for, let us say, the period 1942 to 1955 and plots these data on a graph, measuring quantities bought (in millions of bushels) on the horizontal axis and the price (in dollars) on the vertical axis. The result is the collection of points plotted in Figure 1–4. The points have been labeled with numbers to indicate the years, and the student should note that the points do not follow any particular *chronological* sequence on the graph.

An examination of the points indicates clearly that the quantity consumed tends to be lower the higher the price and that this relationship is rather close and appears to follow a straight line rather well. The investigator "fits"

---

50 cents) in price; thus, the slope of the supply curve is $+1/2$. That is, the absolute value of the slope of the supply curve ($1/2$) is greater than the absolute value of the slope of the demand curve ($1/3$).

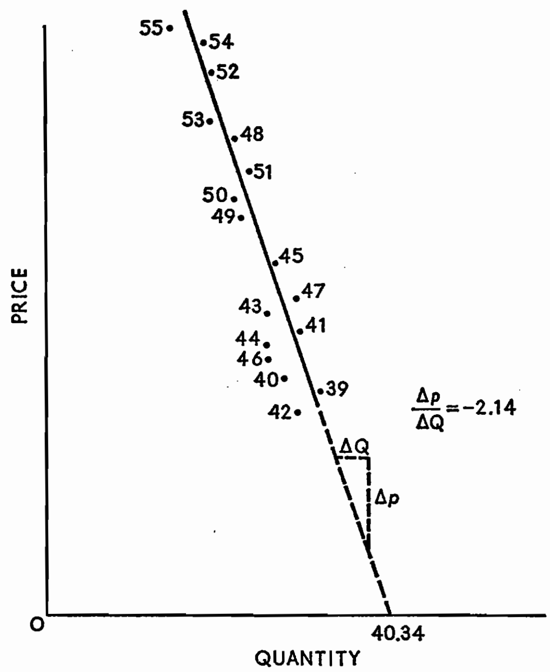
[9] Obviously, such explosive oscillations could not continue indefinitely; some additional factor, such as complications of the sort referred to in footnote 7 above, would eventually bring them to a halt.

a *regression line* to the points by statistical methods.[10] The equation of the best fitting straight line in this particular case is assumed to be

$$Q = 40.34 - 2.14p$$

where Q is the quantity consumed (in millions of bushels) and p is the price (in dollars). That is, the line of the demand curve intersects the quantity axis at 40.34 million bushels, and the slope of the line is —2.14. Starting

FIGURE 1–4
Hypothetical Econometric Demand Curve



from 40.34 million bushels at a price of zero, the quantity consumed declines by about 2.14 million bushels for each rise of one dollar in the price.[11] Thus if the price were to be $2 in, say, 1957 and if no major shift in

[10] It is obviously possible to draw an infinite number of straight lines through the cluster of points. It is therefore necessary to have some criterion for picking the "best" line to draw. The one ordinarily employed (for statistical reasons) is the so-called "least squares" criterion. Thus, the investigator might select the line that would make the sum of the squares of the horizontal deviations of the points from the line as small as possible.

[11] The interpretation of the intercept with an axis of a straight line often proves to be a puzzle to students of economics. The 40.34 million bushels cannot be interpreted as the amount that would be consumed at a price of zero—in fact, it cannot be given a meaningful *economic* interpretation. The line which is fitted to the points is meaningful only

the relationship were to occur, one might estimate that consumption in 1957 would be approximately 36.06 million bushels (40.34 − 4.28). Of course, since the line does not "fit" the points exactly, one would be surprised if consumption at a price of $2 turned out to be exactly 36.06 million bushels, but this would be the best estimate of consumption that one could make on the basis of this model. The existence of deviations of the points from the line indicates that there are factors other than price affecting consumption. It might be possible to improve the ability to predict consumption by taking some of these other factors into account, although it is apparent in this case that the deviations of the points from the line are not very large and that a reasonably accurate prediction of consumption should be possible based on price alone.

In practice, econometric work is a good deal more complex than the above oversimplified hypothetical example would make it appear. A few of the major complications may be indicated. In the first place, most economic relationships involve more than two variables. Thus, it is most unlikely that a satisfactory demand curve for most goods or services could be uncovered without taking account of other variables besides the price. Major variables that might need to be taken into account include consumer income, population, and prices of related goods. However, the necessity of taking account of additional variables causes little difficulty; regression analysis of the type referred to above can readily be extended to multivariable relationships. Another difficulty that is sometimes serious—and, in fact, would be likely to cause trouble in the case discussed above—is the problem of "identifying" the relationship that is being studied. In estimating a demand curve, the identification problem arises due to the fact that there are ordinarily two relationships between the quantity of a good and its price—one on the demand side of the market and the other on the supply side. Unless precautions of some kind are taken, the investigator cannot be sure whether the relationship he uncovers between price and quantity is a demand curve, a supply curve, or merely the historical path traced out by price and quantity as a result of changes in both demand and supply. This problem comes up in almost any extensive econometric investigation. There are various ways of dealing with it, but it does add to the difficulties of the econometrician. Another problem, sometimes serious, is that the data available for use in econometric investigations are subject to errors, often large but not amenable to any kind of measurement.

within approximately the range of values of the variables covered by the points themselves. At a price of zero the commodity would become a free good and no economizing in its use would be necessary. However, when the line fitted to the points is extended to intersect the quantity axis, the intersection occurs at 40.34 million bushels. In order to specify a straight line, it is necessary to employ two constants, one to indicate the level at which the line is to be drawn and the other to specify its slope. The line in Figure 1–4 is specified by indicating (*a*) its intercept with the horizontal axis (40.34), which determines the level of the line, and (*b*) its slope (−2.14). Thus, the intercept has no *economic* meaning but is merely a convenient *mathematical* way to specify the line.

The hypothetical example discussed above is in the field of microeconomics, i.e., the study of the market for a particular good. However, in recent years a good deal of econometric work has been done in the field of macroeconomics. Some studies have attempted to estimate quantitatively particular macroeconomic relationships; for example, there have been numerous quantitative investigations of the forces that determine the aggregate level of consumer spending in the economy.[12] Some studies have gone beyond this and attempted to construct and quantify extensive models of the economy as a whole involving many relationships and potentially capable of predicting the behavior of such basic macroeconomic variables as national product and income, the level of prices, employment and unemployment, and so on.

In this book we shall confine ourselves in the main to the construction of theoretical models with no effort at quantitative investigation. However, we will at many points and wherever it is feasible attempt to draw upon the results of existing econometric studies of the relationships we are investigating. Although these studies are frequently inconclusive and occasionally contradictory, they will at least help us in some cases to arrive at a reasonable judgment as to whether a certain key relationship is of such a magnitude as to be of major importance. In the present state of economic science, about all that can be said of a quantitative nature with any degree of finality relates to general orders of magnitude of various relationships.

OUTLINE OF THE BOOK

In this book we shall be dealing for the most part with large aggregates, such as national income, industrial production, total employment and unemployment, or with concepts that are essentially averages, such as the price level and the wage level. These aggregates and averages must be measured and quantified if they are to be meaningful, and measurement of such concepts involves some fairly complex problems—including some problems to which there are no fully satisfactory answers. Accordingly, Part I of this book deals with the problems of defining and measuring the major macroeconomic variables.

In Part II we develop the theory of income determination. The underlying idea of this section is that production is carried out to satisfy demand so that the level of aggregate demand therefore determines production and income. The determinants of aggregate consumption and investment demand are discussed, as is the role of money and interest rates in determining income.

In the interest of simplicity, the assumption is retained throughout Part II that the economy has sufficient unutilized resources so that changes in aggregate demand have their full impact on real output and employment

[12] Some of these investigations are discussed or referred to in Chapter 5.

with no changes in wages and prices. This obvious defect in the analysis is corrected in Part III, which takes up wages and prices and their relation to employment and unemployment. Chapter 15 in that part deals with the mechanism stressed by the pre-Keynesian classical economists, through which flexible wages and prices were supposed to keep the economy operating at full employment without government intervention. The defects in this mechanism, which underlie the need for monetary and fiscal policy in the modern world, are explained. Chapter 16 discusses inflation and its relation to unemployment in a more realistic setting—a relation which is the source of much difficulty and controversy today.

Parts II and III largely abstract from the complications of economic growth. This additional dimension of macroeconomics is introduced in Part IV. First, the relation between unemployment and inflation developed in Part III is related to the growth path of the economy, and the problems of keeping the economy expanding along this path are discussed. Then the determinants of the growth path itself are taken up, as is the role of government in promoting economic growth.

Throughout Parts II, III, and IV, the analysis is restricted to a "closed economy"—that is, an economy which does not enjoy trade and financial relations with other economies. Part IV "opens up" the economy and shows how problems of foreign trade and the balance of payments complicate the determination of income and employment and raise additional problems for those responsible for the formulation and implementation of monetary and fiscal policy. Since the way in which foreign trade and finance impinge on the domestic economy depends upon the kind of international financial arrangements that prevail, it is necessary to explain at some length the operation of our present international monetary system. We also consider the defects in the present system and some of the reforms that have been proposed to improve it.

Numerous references to economic policy are scattered throughout the book. However, there is no final unified treatment of policy that takes full account of all of the complications that are successively introduced. The reason for this lack is primarily practical: the book is already long, and, in the author's opinion, a fully adequate treatment of policy issues would require another book of at least equal length. It is the author's hope, however, that the careful student of this book will find his knowledge of important issues of public policy substantially expanded and sharpened.[13]

[13] The student who wishes to pursue additional reading more specifically related to policy may find some useful material in Warren L. Smith and Ronald L. Teigen (eds.), *Readings in Money, National Income, and Stabilization Policy* (rev. ed.; Homewood, Ill.: Richard D. Irwin, Inc., 1970). The introduction to Chapter 1 of that book is, in some respects, an extension of this book with a somewhat more specific orientation toward policy. And Chapter 7 (including the introduction) contains a discussion of the problem of using the several instruments of economic policy in a coordinated way to achieve several economic goals.

PART I

Measurement of Macroeconomic Variables

INTRODUCTION

In recent years, a great deal of progress has been made in collecting and disseminating data which may be used to describe, evaluate, and analyze the performance of the economy. The most important body of such data is incorporated into the format of the national income and product accounts. These accounts are explained in Chapters 2 and 3. This material is important for the student to understand, because the national income and product account concepts are frequently used in the remainder of this book.

Evaluation of the economy's performance requires the use of other data that are not included in the national income and product accounts. In particular, there is much concern about employment and unemployment and about the behavior of the price level. In Chapter 3, the problems of measuring employment and unemployment and changes in the price level and the statistical concepts and data that are available for this purpose are explained and evaluated. The chapter also refers to other statistical series that are useful in evaluating the economy's performance.

Chapter 2 THE NATIONAL INCOME AND PRODUCT ACCOUNTS: I

National income and product statistics undoubtedly constitute the single most important collection of economic information available today. These statistics are widely used by government agencies in their evaluation of the performance of the economy and in arriving at decisions concerning economic policy, as well as by businessmen in judging economic trends and in studies of the markets for their products. The systematic collection and presentation of these statistics and their progressive refinement represents a major forward step in the scientific study of economic phenomena in the last 25 years or so, not only in the United States but in most other countries as well.[1]

At the outset of the discussion it is important to stress the systematic nature of these statistics. As we shall see, they are compiled and presented within an *accounting framework.* In fact, *national income accounting* is one branch of the subject of *social accounting,* which has undergone an extensive development in recent years.[2] Thus while arbitrary decisions must be made concerning the handling of certain items, the national income accounts

[1] The official compilation of national income and product statistics in the United States is carried out by the Office of Business Economics of the Department of Commerce. The statistics appear regularly in *The Survey of Current Business,* published monthly by the Department of Commerce. The July issue of the *Survey* each year contains an extensive compilation. Summary tables of national income statistics are to be found in the *Federal Reserve Bulletin,* published monthly by the Board of Governors of the Federal Reserve System, and tabulations are contained in the *Economic Report of the President,* which is issued in January of each year. A lengthy discussion of methodology, conceptual problems, and sources of data, together with statistics for past years, is contained in *National Income: A Supplement to the Survey of Current Business,* 1954 edition (Department of Commerce, 1954). Further information and some improvements in statistical formulation are to be found in *U.S. Income and Output: A Supplement to the Survey of Current Business* (Department of Commerce, 1958). Recent conceptual changes are discussed in the *Survey of Current Business,* August 1965, pp. 6–28. Revised estimates for earlier years are to be found in *The National Income and Product Accounts of the United States, 1929–1965 Statistical Tables* (Department of Commerce, 1966).

[2] Other branches of social accounting include balance-of-payments accounting, money flows accounting, and input-output accounting. Balance of payments accounting is discussed briefly in Chapter 9. For a more extensive discussion of social accounting, see John P. Powelson, *Economic Accounting* (New York: McGraw-Hill Book Co., 1955).

possess an analytical foundation, and it is important for the student to grasp the logic of the subject.

The national income or product may be viewed in three different ways, all of which are important and all of which will be employed in our discussion. First, the total output or product of the economy may be looked upon as the total amount spent on final output by all the participants in the economic process, including consumers or households, business concerns, governmental units, and foreigners. Second, the total national income or output may be viewed as the sum of all the incomes earned by participants in the economic process for productive contributions made either by themselves or by their property. Finally total product may be construed as the sum of the output produced by all of the producing units in the economy.

GROSS NATIONAL PRODUCT AND EXPENDITURES

Gross national product is defined as the total gross value, at market prices, of the output of all goods and services produced in the economy during a particular period. It is "gross" in the sense that no deduction is made for the amount of existing capital—plants, equipment, tools, etc.—used up through wear and tear during the period. However, it is "net" in the sense that repeated resales of the same goods have been eliminated.

GNP may be looked upon as the total amount spent on final output by all economic units. For the year 1968, GNP as the sum of expenditures breaks down as follows (in billions):

Personal consumption expenditures	$536.6
Gross private domestic investment	126.3
Net exports of goods and services	2.5
Government purchases of goods and services	200.3
Gross National Product	$865.7

A more detailed breakdown, to which we shall make frequent reference, is given in Table 2–1.[3]

The economy is divided into four sectors for purposes of national income accounting. These four sectors are households, business firms, rest of the world (or the foreign sector), and government. The GNP, as shown in the table, is the summation of amounts spent on final output by these four sectors. We shall consider briefly the expenditures of each of them.

1. The Household Sector. The expenditures of the household sector are called *personal consumption expenditures.* Such expenditures totaled $536.6 billion in 1968 and, as shown in Table 2–1, this amount can be broken down into $83.3 billion for durable goods, $230.6 billion for nondurable goods, and $222.8 billion for services. Durable goods include all

[3] All of the national income and product data for 1968 presented in this chapter and the next are taken from "U.S. National Income and Product Accounts, 1965–68," *Survey of Current Business,* July 1969, pp. 13–50.

goods bought by households having a normal life of more than one year, with the exception of new houses, which are included under the heading of investment rather than consumption. Thus, durable goods include new automobiles, radio and TV sets, electric appliances, furniture, and the like. Nondurable goods include tangible items, such as food and clothing, having a presumptive life of less than one year. Services include such intangible items as recreational expenditures and expenditures on medical care and education.[4]

TABLE 2–1

Gross National Product or Expenditure, 1968
(in billions)

Personal consumption expenditures				$536.6
Durable goods			$ 83.3	
Nondurable goods			230.6	
Services			222.8	
Gross private domestic investment				126.3
Fixed investment			$119.0	
Nonresidential		$88.8		
Structures	$29.3			
Producers' durable equipment	59.5			
Residential structures		30.2		
Change in business inventories			7.3	
Net exports of goods and services				2.5
Exports			$ 50.6	
Imports			48.1	
Government purchases of goods and services				200.3
Federal			$ 99.5	
State and local			100.7	
Gross National Product				$865.7

NOTE: Items may not add to totals due to rounding.
SOURCE: Department of Commerce.

The national income accounting system contains only four sectors, and it is, of course, necessary that every economic unit be assigned to one of them. As a result of this, some units are assigned to the household sector for lack of a better place to put them, in spite of the fact that they do not fit very well. Thus, charitable and nonprofit institutions and private pension and welfare funds are included in the household sector, and, accordingly, the expenditures on goods and services of these organizations are included in personal consumption expenditures. On the other hand, most transactions between consumers and such institutions are eliminated from the accounts entirely, as are gifts between households proper, since they cancel out in the

[4] The Department of Commerce prepares supporting tables giving estimates of personal consumption expenditures broken down in considerable detail by type of product. See *The National Income and Product Accounts of the United States, 1929–1965 Statistical Tables, op. cit.*, pp. 40–50.

process of consolidating the components of the household sector. However, some transactions within the household sector are included since they represent payments for services rendered; the main item here is payments for domestic service.[5]

2. The Business Sector. Gross private domestic investment represents the gross expenditures on final output by the business sector, which takes in all private profit-oriented economic units, including both corporations and unincorporated enterprises. Gross private domestic investment can be divided into two major components: (1) *gross expenditures on fixed capital* and (2) *change in business inventories.*[6] Total gross private domestic investment in 1968 amounted to $126.3 billion, which can be broken down (see Table 2–1) into fixed capital investment of $119 billion and an increase in business inventories of $7.3 billion.

Expenditures on *new residential construction* are included, along with expenditures on industrial, commercial, and agricultural construction, in gross private domestic investment. Thus, housing expenditures are treated differently from other household expenditures.[7] In 1968, residential construction expenditures amounted to $30.2 billion, as shown in Table 2–1. Other construction, which includes predominantly business expenditures on the construction of plants, stores, and so on, amounted to $29.3 billion.[8] The remainder of fixed investment was accounted for by $59.5 billion of expenditures on producers' durable equipment, which includes purchases of machinery, tools, store counters, and the like by business concerns. All of these expenditures on fixed capital goods are *gross* in the sense that they include both expenditures on *new* plants, equipment, and houses, which add to the total existing supply of such assets, and *replacement* expenditures, which merely make good destruction (by fire, flood, etc.) and depreciation (wear and tear) on existing assets.

The remainder of gross private domestic investment in 1968 is accounted for by an increase of $7.3 billion in business inventories. This represents the net increase during the year in stocks of goods—including finished goods, partly finished products, and raw materials—by all business concerns, incorporated as well as unincorporated.[9] The student should note carefully that

[5] Thus, there is a small amount of income produced or generated *within* the household sector.

[6] The student should note that "investment" in our terminology means expenditures on real capital—plants, equipments, inventories, etc.—by business. Purchases of bonds and stocks are not investment, although, of course, they may indirectly facilitate real investment by providing funds to business firms desiring to engage in such activity.

[7] Home ownership—even in cases where the house is occupied by its owner—is treated as a business activity. For a fuller discussion, see Chapter 3, pp. 57–59.

[8] Government construction—roads, dams, schools, other public buildings—is included in government purchases of goods and services.

[9] The change in the value of inventories as shown by the accounts of businesses is corrected by an "inventory valuation adjustment." In 1966, the change in inventories as shown by businessmen's books was an increase of $15.4 billion, but the inventory valua-

the concept of inventory investment relates to the *change* in inventory holdings during the period and not to the total *stock* of inventories in warehouses or on store shelves at the end of the period. Thus, inventory investment can be, and often is, negative; that is, inventory stocks may be reduced as well as increased.[10]

Purchases of goods (or services) by one business firm from another are not included in GNP expenditures except to the extent that they (*a*) represent purchases of fixed capital goods having a life of more than one year, or (*b*) result in an increase in inventory stocks. Otherwise purchases by one business are netted against sales by others, so that the amounts disappear in the process of consolidating the accounts of the business sector. If interfirm transactions were not "netted out" in this way, they would result in double counting; for example, the steel which is contained in automobiles would be counted once as part of the product of the steel industry and again as part of the product of the automobile industry.

*3. **The Rest-of-the-World Sector.*** This sector includes all economic units (households, businesses, and governments) which are not residents of the United States. Table 2–1 shows that in 1968 *net exports* to the rest-of-the-world sector amounted to $2.5 billion. This amount represents the difference between total exports of goods and services of $50.6 billion and total imports of goods and services of $48.1 billion. It should be noted that the amount of net exports of goods and services can be, and on occasion is, negative.

The discerning student may wonder why *net* exports of goods and services is one of the components of the *gross* national product. The explanation of this is quite simple. If we knew the purchases of *domestically produced* final output by the four sectors, households, business, rest-of-the-world, and government, we could total these figures to get the gross national product. However, we import large quantities of goods and services, and these imports—either in their original form or after further processing—may go to households in the form of consumer goods, go to business in the form of fixed capital goods or inventories, be reexported to foreign countries, or be included in goods purchased by domestic governmental units. It is impossible to track down these imports; thus our data on expenditures by households, businesses, rest-of-the-world, and government relate to *total* rather than merely *domestically produced* goods and services. Since all imports must be distributed to one of these four sectors, when these expendi-

---

tion adjustment was negative to the extent of $2 billion, thus reducing the increase in inventories as shown in the national accounts to $13.4 billion. The logic of the inventory valuation adjustment is explained in Chapter 3, pp. 56–57.

[10] *Gross fixed* investment cannot be negative, since it represents the total value of fixed capital goods produced, and production is inherently nonnegative. However, gross private domestic investment can be negative if inventory investment is negative and exceeds gross fixed investment. And *net fixed* investment can be negative if gross fixed investment (i.e., production of fixed capital goods) is less than depreciation.

tures are totaled the result is gross national product *plus imports.* Thus, a correction in the form of the deduction of imports is required. In getting the gross national product as the sum of expenditures, this is handled by deducting imports from exports to get net exports, which is included as an expenditure component.

4. ***The Government Sector.*** The final component of GNP expenditures is *government purchases of goods and services,* which amounted to $200.3 billion in 1968. Federal government purchases were $99.5 billion, while purchases of state and local government units accounted for the remaining $100.7 billion, as shown in Table 2–1. The fact that these are government expenditures for the purchase of *goods and services* from the household, business, and rest-of-the-world sectors should be stressed. As we shall see, there are other government expenditures which are not included in the GNP.

NATIONAL INCOME BY TYPE OF INCOME

We shall now shift the focus of our study briefly from the gross national product to the national income, then turn our attention to the relation between the two concepts. The *national income* may be defined as total income *earned* by participants in production through productive contributions made by them or by their property during a particular period. National income may be shown as a summation of incomes earned in various forms. For 1968, the tabulation is as follows (in billions):

Compensation of employees	$513.6
Proprietors' income	63.8
Rental income of persons	21.2
Corporate profits and inventory valuation adjustment	87.9
Net interest	28.0
National Income	$714.4

A more detailed breakdown, to which we shall make reference, is given in Table 2–2.

Compensation of employees includes all wages and salaries earned by households, together with other elements of labor income, including such "fringe benefits" as contributions made by employers to government and private pension and welfare funds on behalf of their employees (see breakdown in Table 2–2). *Proprietors' income* includes the net profits of unincorporated business enterprises. As indicated in Table 2–2, this covers incomes of nonfarm unincorporated businesses and self-employed professional people (doctors, lawyers, teachers, accountants, etc.), and incomes of farm operators. *Rental income of persons* includes incomes in the form of rent and royalties received from the ownership of property. *Corporate profits* includes the profits of all private corporations; Table 2–2 shows the portion of such profits paid to the government in the form of corporate income taxes, the portion paid to stockholders in dividends, and the portion (undistributed profits) retained by the corporations themselves to finance expansion of their operations. *Net interest* includes interest earned by households and governments minus interest paid by households and governments. That is, interest paid by households and governments is excluded from the national income. The reasons for this exclusion are explained later.[11] The inventory

TABLE 2–2

National Income by Type of Income, 1968
(in billions)

Compensation of employees			$513.6
Wages and salaries		$465.0	
Supplements to wages and salaries		48.6	
Employer contributions for social insurance	$24.4		
Employer contributions to private pension and welfare funds	20.1		
Other supplements	4.0		
Proprietors' income			63.8
Business and professional		$ 49.2	
Income of unincorporated enterprises	$49.9		
Inventory valuation adjustment	−0.7		
Farm		14.6	
Rental income of persons			21.2
Corporate profits and inventory valuation adjustment			87.9
Profits before tax		$ 91.1	
Profits tax liability	$41.3		
Dividends	23.1		
Undistributed	26.7		
Inventory valuation adjustment		−3.2	
Net interest			28.0
National Income			$714.4

NOTE: Items may not add to totals due to rounding.
SOURCE: Department of Commerce.

valuation adjustments which are included with business and professional income (see Table 2–2) and with corporate profits will also be explained later.[12] All incomes are *before deduction of personal taxes and corporate income taxes.*

RELATION BETWEEN GNP AND NATIONAL INCOME

Having developed the GNP as the sum of expenditures on final output by the four sectors and the national income as the sum of the incomes of various types earned by the participants in production, let us now consider how these two concepts are related to each other. The relationship for 1968 is summarized in the following table (in billions):

[11] See Chapter 3, pp. 43–45.

[12] See Chapter 3, pp. 56–57, below; also footnote 9 above.

Gross National Product	$865.7
Less: Capital consumption allowances	73.3
Equals: Net National Product	$792.4
Less: Indirect business tax and nontax liability	77.9
Business transfer payments	3.4
Statistical discrepancy	−2.5
Plus: Subsidies minus current surplus of government enterprises	0.8
Equals: National Income	$714.4

Capital consumption allowances consist mainly of allowances for the depreciation of capital through normal wear and tear or obselescence, but they also include allowances for the destruction of capital through such accidents or catastrophes as fires and floods.

When capital consumption allowances are deducted from GNP, we obtain the *net national product.* This may be defined as the net value of the total output of the economy valued at market prices for the period. We could, of course, deduct capital consumption allowances from gross private domestic investment, thus obtaining an amount which we might call "net private domestic investment." Then we could present the net national product as the sum of personal consumption expenditures, "net private domestic investment," net exports, and government purchases of goods and services, in a manner analogous to our presentation of gross national product from the expenditure side. However, this interpretation is open to serious objections. The measurement of net investment is extremely difficult for reasons that are partly theoretical and partly due to the inadequacies of the statistical data. The conceptual problem may be illustrated by noting the fact that if an industrial plant is destroyed, wears out, or becomes obsolescent, so that the firm is faced with the necessity of replacing it, the plant will be replaced not with one exactly like it but with one which incorporates the latest techniques and is probably more productive than the original plant. The question is: how much of the expenditure for the new plant represents "replacement" and how much represents "new investment" or "expansion"? This question is extremely difficult to answer in principle, and it is virtually certain that the accounting allowance made by the firm to cover the destruction or wear and tear on the old plant will not provide a satisfactory estimate of the amount of replacement. In addition to the basic conceptual difficulty in drawing a satisfactory line between replacement and new investment, the value of the accounting allowances is impaired by several purely statistical or accounting difficulties. Depreciation (which constitutes by far the greatest portion of capital allowances) is ordinarily computed according to some rather simple rule and is based upon the original cost of the asset.[13]

[13] Straight-line depreciation is the simplest and is very commonly used. Under this procedure, for example, if a plant costs $200,000 to build and is expected to last for 40 years, at the end of which time its scrap value will be negligible, $5,000 of depreciation will be charged to expenses of operation each year.

Since the assets being depreciated were purchased in many different years at prices which varied from year to year, while replacement must be made at the prices prevailing in the current year, the depreciation charge is likely to bear little relation to the actual deterioration of capital during the year.[14] A further difficulty is that methods used in computing depreciation are strongly influenced by the provisions of the tax laws, which are changed from time to time.

The problem of measuring net investment is an important one in economics, since net investment is one of the basic sources of economic growth, and some reasonably respectable estimates of net investment have been made. But it should be apparent from a consideration of the difficulties sketched above that the estimate of net investment derived by subtracting capital consumption allowances from gross private domestic investment as shown in the national income accounts is of little value. As a matter of fact, this difficulty is serious enough to raise doubts about the validity of the estimates of net national product, and in practice little use is made of this concept.

Having deducted capital consumption allowances from GNP to get net national product, the next step in getting to national income is to deduct *indirect business tax and nontax liabilities.* This item includes all taxes levied on business except corporate income taxes, together with fees paid by business to government units for specific services. All sales taxes, whether ostensibly levied on businesses or on consumers, are included—in fact, sales taxes, together with taxes on business property, make up the major portion of the item.

Thus, indirect business taxes are excluded from national income—that is, they are deducted from gross national product in the process of arriving at national income; whereas all other types of taxes—corporate profit taxes, personal tax and nontax payments, and social security contributions—are included in national income. Partly this reflects the fact that indirect business taxes are, in general, treated as expenses and deducted from revenues in the process of computing income by the business units that pay them. However, it also reflects an assumption on the part of the national income accountants concerning the incidence of taxes: an assumption that all taxes of the indirect business type are passed on to consumers in the form of higher product prices. Undoubtedly this assumption is not wholly correct; some sales taxes, for example, are surely borne by the profits of the firms that pay them or are shifted backward in the form of reduced incomes to the factors of production which produce the firms' products. But in view of the difficulties involved in judging the incidence of specific taxes, the procedures adopted by those who have constructed the national income accounts are

[14] The problem of proper accounting for depreciation in an economy characterized by a variable price level is one which has received a great deal of attention among accountants in recent years, although no satisfactory solution has yet been found.

probably the most "neutral" and generally reasonable ones that could be applied.

In addition to the deduction of indirect business taxes, two other relatively minor adjustments are made in getting from net national product to national income. One is the deduction of *business transfer payments.* These are payments made by the business sector to the household sector for which no services are rendered to the business sector. An example is a contribution by a corporation to a charitable institution, such as a community fund. This is a payment to the household sector, since, as we have seen, charitable institutions are included in that sector. It is not eligible for inclusion in the national income, since only payments made for services rendered in producing current output should be so included. It is a charge against the gross national product, however; that is, it must be covered out of the receipts of the business from the sale of its product. For these reasons, it is deducted from the net national product in the process of arriving at the national income. A similar example is losses suffered by business due to theft or as a result of defaults on debts owed to business by households, since, in effect, to the extent of these losses the business sector has made gifts of goods and services to the household sector.

The other adjustment that must be made in deriving the national income from the net national product is the addition of *subsidies minus current surplus of government enterprises.* Subsidies are payments made by the government sector to the business sector for which the government sector receives no goods or services in return.[15] They constitute a source in addition to sales proceeds from which business concerns may make factor payments or derive additional profits. For this reason, they must be added to net national product to obtain national income. Current surpluses—or profits—of government enterprises, such as municipally owned electric power plants, are deducted in arriving at national income because they are derived from the sales proceeds of such enterprises but are not distributed as income to anyone.[16]

Business transfer payments and subsidies minus current surplus of government enterprises are quite unimportant items in the national income accounts; moreover subsidies may be regarded as merely negative indirect

[15] When the government purchases products at prices higher than would prevail in a free market in order to maintain the income of their producers, such payments are technically not subsidies but purchases of goods and services and are so classified in the national income accounts. This is the case with respect to some farm price support payments. On the other hand, to the extent that these payments take the form of direct cash grants to farmers or result in losses on nonrecourse loans by the Commodity Credit Corporation (a government enterprise) they are reflected in subsidies minus current surplus of government enterprises.

[16] The fact that some government enterprises pay subsidies—e.g., the postal system, which is treated as a government enterprise, pays subsidies in connection with air mail contracts—makes it difficult to separate entirely subsidies and profits of government enterprises. Hence, the two are incorporated in one item in the accounts.

business taxes. Consequently, the student may properly think of national income as being equal to net national product minus indirect business taxes. This brings us to an alternative definition of the national income that is sometimes used: national income is the total net product of the economy valued at *factor costs.* Thus, national income is contrasted with net national product, which is the net product valued at *market prices.* That is, national income is the net output valued at what was paid to the factors of production for producing it, whereas the net national product is the same output valued at what it sold for on the market, the difference between the two being indirect taxes which are assumed to be added to costs in the determination of prices.[17]

The methods used by the Department of Commerce in preparing estimates of national income and product may be outlined very briefly as follows. First, estimates are prepared of expenditures by the four sectors—that is, personal consumption expenditures, gross private domestic investment, net exports, and government purchases of goods and services. This is, of course, a very difficult task, and the Department makes use of every scrap of information it can obtain from either governmental or private sources. As we have seen, the sum of these categories of expenditures provides an estimate of the gross national product. Next, estimates are prepared of capital consumption allowances, indirect business taxes, business transfer payments, and subsidies minus current surplus of government enterprises. When the first three of these items are subtracted from and the last item is added to the estimate of the gross national product, the result is an estimate of the national income. Another estimate of the national income is obtained by estimating the various types of income—compensation of employees, proprietors' income, rental income of persons, corporate profits, and net interest—and adding them together. The *statistical discrepancy* is the difference between the two estimates of national income. The estimate of national income derived from the income side is taken to be the "true" estimate, and if the estimate of national income derived through the expenditure side is less than this amount, there is a negative statistical discrepancy which is *subtracted* from the estimate derived from the expenditure side to make the two estimates balance.[18] For example, in 1968 the estimate derived from the expenditure side was less than the estimate derived from the income side by $2.5 billion, so the statistical discrepancy was negative in that amount. If the

[17] It should be noted that when it is said that the national income is net output valued at factor costs, the word "cost" is being used in a rather peculiar way, since *all* payments to factors (including profits) are here regarded as costs, whereas cost, to the economist, usually means only those payments that are necessary to induce the performance of productive services.

[18] Of course, it could equally well be said that the statistical discrepancy is the difference between two estimates of GNP, one derived from the expenditure side and taken to be the "true" estimate and the other derived from the income side.

estimate derived from the income side is less than that derived from the expenditure side, the discrepancy will, of course, be positive.

The statistical discrepancy is an estimate of the *consistency* but not the *accuracy* of the estimates. That is, there may be offsetting errors in various items which will not be reflected in the statistical discrepancy.[19] As far as the accuracy of the estimates is concerned, a general evaluation is impossible. Some items can be estimated with a high degree of accuracy with the data available, while for other items the data are poor and the estimates not very reliable. Users of the statistics should study the estimating techniques and sources of data used by the Department of Commerce and form their own judgments concerning the reliability of the estimates of the particular concepts they are interested in.[20]

HYPOTHETICAL ILLUSTRATION OF BUSINESS ACCOUNTS

At the outset of our discussion, it was pointed out that the national income or product could be viewed as (1) the sum of the amounts spent on final output by the various groups or sectors of the economy, or (2) the sum of all incomes earned from productive activity, or (3) the sum of the outputs of all producing units in the economy. In our discussion up to this point, we have developed and made use of the first two approaches, but no attention has been given to the third. We now turn our attention to this third approach.

Table 2–3 gives the income statement for the year 1968 of a hypothetical corporation which we shall call the Excelsior Electronics Corporation. No particular effort has been made to make the figures or accounts included in this statement realistic, and some of the items are included mainly because of their value in illustrating the concepts of national income accounting. The income statement of a firm—sometimes called the profit and loss statement —summarizes the firm's operations during a specified period of time. It begins with receipts from the sale of products ($32,347,000 in our example), and deducts the cost of goods sold ($14,291,000), thus giving gross profit from sales ($18,056,000). The cost of goods sold is computed by taking the purchases of materials ($16,127,000) and adding the reduction in inventories or deducting the increase in inventories as the case may be. In

---

[19] For example, if personal consumption expenditure was overestimated by $2.1 billion and at the same time compensation of employees was overestimated by $2.2 billion, with everything else estimated accurately, the estimate of national income from the expenditure side would be overstated by $2.1 billion, the estimate derived from the income side would be overestimated by $2.2 billion, but the statistical discrepancy would be only a negative $0.1 billion.

[20] Sources of data and methods of estimating are explained at considerable length in *National Income: A Supplement to the Survey of Current Business, op. cit.,* Part III. See also *U.S. Income and Output: A Supplement to the Survey of Current Business, op. cit.,* chap. 7.

TABLE 2–3

Hypothetical Income Statement
EXCELSIOR ELECTRONICS CORPORATION
Calendar Year 1968
(in thousands)

Item Number	*Name of Item*		
1	Sales		$32,347
2	Less: Cost of goods sold		14,291
3	Purchase of materials	$16,127	
4	Plus: Inventory, Jan. 1, 1968	5,382	
5	Less: Inventory, Dec. 31, 1968	7,218	
6	Gross profit from sales		$18,056
7	Less: Operating expenses		13,473
8	Wages and salaries	$ 6,124	
9	Commissions and bonuses	721	
10	Traveling expenses	497	
11	Office supplies	376	
12	Gas, water, electricity, and fuel	1,845	
13	Telephone and telegraph	237	
14	Insurance	131	
15	Rent and royalty payments	1,498	
16	Charitable contributions	37	
17	Depreciation	1,284	
18	Property taxes	723	
19	Net operating income		$ 4,583
20	Plus: Nonoperating income		220
21	Subsidies from government	$ 188	
22	Interest income	32	
23	Less: Nonoperating expenses		280
24	Interest expense	$ 264	
25	Losses on bad debts	16	
26	Net profit before income taxes		$ 4,523
27	Less: Income taxes accrued		2,169
28	Net profit to stockholders		$ 2,354
29	Less: Dividends		1,494
30	Profits retained in the business		$ 860

our example, inventories increased by $1,836,000 during the year, since they were $5,382,000 at the beginning of the year and $7,218,000 at the end. Thus, the company, in effect, used up in production only $14,291,000 of materials (the difference between total materials purchased and the increase in its inventories). Total operating expenses, ($13,473,000, as listed in Table 2–3) are then deducted from gross profit from sales to obtain net operating income ($4,583,000). Adding nonoperating income ($220,000) and subtracting nonoperating expenses ($280,000), we obtain net profit before income taxes ($4,523,000). This is divided between income taxes ($2,169,000), dividends paid to stockholders ($1,494,000), and profits retained in the business to finance expansion ($860,000).

It is possible to take the information contained in the income statement and reorganize it in such a way as to bring out the firm's contributions to national product and income. Such a reorganization for the Excelsior Electronics Corporation is presented in Table 2–4, which contains exactly the same information as Table 2–3 organized in a different way.

Looking first at the right-hand side of Table 2–4, which shows the sources of our company's productive contribution, we note that the company purchased materials from other firms in the amount of $16,127,000. In addition, a number of its other expense items are assumed to represent purchases of goods or services from other firms, including $497,000 for traveling expenses, $376,000 for office supplies, $1,845,000 for gas, water, electricity, and fuel, $237,000 for telephone and telegraph service, and $131,000 for insurance. Thus, altogether it purchased goods and services amounting to $19,213,000 from other firms. Applying the labor services of its employees, the use of its equipment, and its technical skill, Excelsior Electronics, in effect, processed these purchased goods and services into products which it was able to sell for $32,347,000 and in addition it increased its inventories by $1,836,000, for a total value of $34,183,000. Thus, it added $14,970,000 ($34,183,000 minus $19,213,000) to the value of the materials and services it purchased from other firms. This amount of *value added* through productive activity is a measure of the value of the firm's production. If we were to compute the value added of each economic unit in which productive activity is carried on and add these amounts together, we would obtain the GNP. Thus, we may say that this amount of $14,970,000 is the *gross product of the Excelsior Electronics Corporation or its contribution to GNP.*

If we were to consolidate accounts corresponding to the sources side of Table 2–4 for all business firms in the economy, we would find that in the consolidation, the amounts of purchases from other firms would cancel out an equivalent amount of sales (since the purchases of materials by all firms must necessarily be equal to the sales of materials of all firms). Thus, sales and purchases of materials for further processing between firms would both be eliminated and we would be left with only the sales to consumers, to business for fixed capital investment, to foreigners, and to the government, together with the change in inventories. Thus, the consolidation of the sources sides of all the accounts would give us the total product produced by the business sector as the sum of expenditures on this product.

The left-hand or allocations side of Table 2–4 shows how the value created as a result of the activities of the Excelsior Electronics Corporation was distributed or allocated. This side of the table includes all of the remaining items from Table 2–3 that were not used in the construction of the sources side of Table 2–4. Thus, of the $14,970,000 of value created by the company, $1,284,000 was set aside to cover depreciation on its plant, equipment, and tools; $723,000 was paid in property taxes; $37,000 went for charitable contributions; $16,000 to cover losses on bad debts;

TABLE 2–4

Statement of Sources and Allocations
EXCELSIOR ELECTRONICS CORPORATION
Calendar Year 1968
(in thousands)

Item No. from Table 2–3	*Allocations*	
17	Depreciation	$ 1,284
18	Property taxes	723
21	Less: Subsidies from government	−188
16	Charitable contributions	37
25	Losses on bad debts	16
8	Wages and salaries	6,124
9	Commissions and bonuses	721
15	Rent and royalty	1,498
24	Interest expense	264
22	Less: Interest income	−32
26	Net profit before income taxes	4,523
	Total Allocations of Value Added	$14,970

Item No. from Table 2–3	*Sources*		
1	Sales		$32,347
	Increase in inventories		1,836
5	Inventory, 12/31/68	$ 7,218	
4	Less: Inventory 1/1/68	5,382	
	Sales plus increase in inventories		$34,183
	Less: Purchases from other firms		19,213
3	Purchases of materials	$16,127	
10	Traveling expenses	497	
11	Office supplies	376	
12	Gas, water, electricity, and fuel	1,845	
13	Telephone and telegraph	237	
14	Insurance	131	
	Value Added		$14,970

$6,124,000 was paid to employees as wages and salaries; $721,000 was paid as commissions and bonuses; $1,498,000 was paid as rent on property or royalties on patents, etc.; and $264,000 was paid as interest. There are two negative items, subsidies received from the government of $188,000 and interest income of $32,000. These items are entered with negative signs because instead of using up value created by operations, they actually provide supplementary sources of funds to be allocated. The remaining item, representing the difference between total revenue and total expenses, is the

TABLE 2–5

Contributions of Excelsior Electronics Corporation to Various National Product and Income Aggregates, Calendar Year 1968
(in thousands)

Contribution to	*Composition (Item Numbers from Table 2–3)*	*Amount*
Gross National Product	Value added*	$14,970
Less: Capital consumption allowances	17	1,284
Net National Product	...	$13,686
Less: Indirect business taxes	18	723
Business transfer payments	16 + 25	53
Plus: Subsidies minus current surplus of government enterprises	21	188
National Income		$13,098
Compensation of employees	8 + 9	6,845
Proprietors' income	...	....
Rental income of persons	15	1,498
Corporate profits	26	4,523
Net interest	24 − 22	232

* See calculations on the "Sources" side of Table 2–4.

profit of the company before taxes, which amounts to $4,523,000, just sufficient, of course, to make the two sides of the table balance.

In Table 2–5, the material from Tables 2–3 and 2–4 is reorganized into categories corresponding to the items in the national income and product accounts in order to show the contributions of the Excelsior Electronics Corporation to the various national product and income aggregates. Its contribution to GNP is, of course, the amount of its value added, as explained above. Its contribution to capital consumption allowances is its depreciation charges of $1,284,000, its contribution to indirect business taxes is its property taxes of $723,000, and so on. The second column gives the numbers of the items from Tables 2–3 and 2–4 which are contained in each item of Table 2–5. It is apparent that the contribution to national income can be derived in two ways: one, by starting from the contribution to gross national product and making the adjustments already discussed; and the other, by adding together the contributions of the company to the different components of income. There is, of course, no statistical discrepancy shown in Table 2–5, because we are working with accounts that balance exactly, so that the two estimates of national income must necessarily be equal.

The value added of a governmental unit may also be computed. This is done by taking the total purchases of goods and services by the unit and subtracting the amount of its purchases from business. When this is done, the remainder represents the amount paid by the government to its own employees. Thus, the gross product of government in the accounts is merely the value of the services of government employees as measured by the salaries they receive.[21]

Although the calculation of the values added by individual units and the summation of these figures is not a practical procedure for computing the national totals, it is of considerable pedagogical value as a means of bringing out the logic of the national income accounting system. The Department of Commerce prepares estimates of the gross product and national income by industries and sectors of the economy. For example, the estimates for 1968 of the gross product of the four basic sectors are as follows (amounts in billions):

Household sector	$ 25.2
Business sector	740.6
Rest-of-the-world sector	4.7
Government sector	95.2
Gross National Product	$865.7

The product of the household sector is mainly the value of the services of paid domestic employees. The product of the rest-of-the-world sector consists of the pay of permanent United States residents employed in the United States by foreign governments and international organizations, together with the net inflow of interest and dividends on investments by United States residents in foreign countries.

[21] This is true of general government activities. However, government-owned enterprises (e.g., a municipally owned electric power plant) which sell their products or services in the market are handled in the national accounts in a manner very similar to private business enterprises. The handling of general government activities involves a number of problems which are discussed in Chapter 3.

Chapter 3 THE NATIONAL INCOME AND PRODUCT ACCOUNTS: II

In the previous chapter, we developed the three basic approaches to the national accounts and explained the GNP, the national income, and the relation between these concepts. In this chapter, we shall explain two other frequently used income concepts, personal income and disposable personal income, and then discuss certain other aspects of the national income and product accounts, including some of the problems that are encountered in their construction.

PERSONAL INCOME

As we have seen, national income is the total amount of income *earned* as a result of productive services rendered in producing current output. For some purposes, however, we may be interested in the amount of income *received* by households. Income received and income earned are not the same. Some incomes are earned but not currently received, while some incomes are received which are not earned through current productive activity. The amount of income effectively received by the household sector is called *personal income*. Personal income can be derived from national income by (1) adding elements of income received but not earned, and (2) deducting elements of income earned but not received. The following table presents the derivation of personal income from national income for 1968 (amounts in billions):

National Income		$714.4
Less:	Corporate profits tax liability	41.3
	Undistributed corporate profits	26.7
	Corporate inventory valuation adjustment	−3.2
	Contributions for social insurance	47.0
	Wage accruals less disbursements	....
Plus:	Government transfer payments to persons	55.8
	Interest paid by government (net) and by consumers	26.1
	Business transfer payments	3.4
Equals:	Personal Income	$687.9

Corporate profits and inventory valuation adjustment ($87.9 billion) consists of four parts (see Table 2–2, page 31): (1) corporate profits tax liability, $41.3 billion; (2) undistributed corporate profits, $26.7 billion; (3) corporate inventory valuation adjustment, $—3.2 billion; and (4) dividends, $23.1 billion. The entire amount—i.e., all four parts—is included in national income, but the only part that is actually distributed to households is dividends. Consequently, the other three parts must be subtracted out in the process of adjusting national income to get personal income; that is, they represent elements of income earned but not received by households. That is the reason for the first three deductions in the above table.[1]

Contributions for social insurance, which include both employer and employee contributions (payroll taxes) in connection with the federal Old Age and Survivors Insurance program, the federal disability insurance program, and the Medicare program, as well as federal and state unemployment compensation plans, are deducted because they are included in national income as part of employee compensation but are siphoned off directly to the government and not received by the workers.[2]

Turning to the amounts to be added in adjusting national income to get personal income, *government transfer payments to persons* are payments by governmental units to the household sector for which the government receives no current services in return. These include social security benefits —old age, disability, medical care, and unemployment compensation—relief payments, pension payments to retired government employees, veterans' benefits, and so on. Since no current services were rendered in return for these payments, they are not included in national income, but since they do represent payments currently received by households, they are included in personal income.[3]

Interest paid by government (net) and by consumers is not included in national income, since it is not regarded as payment made in return for

---

[1] An alternative way of making the adjustment that is more frequently used is to *deduct* the *entire* amount of corporate profits and inventory valuation adjustment ($87.9 billion) and *add* dividends ($23.1 billion).

[2] It may be noted that it is not necessary to deduct contributions to private pension and welfare funds for the reason that these funds are included, for lack of a better place to put them, in the household sector of the accounts.

Wage accruals less disbursements (which was zero in 1968) is an unimportant item. Not all wages that are *earned* and therefore included in national income need necessarily be *received* during the period. (For example, the period may end on Wednesday while the wages of some employees will not be paid until Friday). Wages earned but not received are deducted from national income to get personal income. The adjustment will be negative if the amount of wages earned but not received declines between the beginning and end of the period, since this means that employers disburse more wages than are earned in that period. However, the adjustment is always small and frequently is zero.

[3] It may be noted that, since private pension funds are included in the household sector, employer contributions and property income (interest, dividends, etc.) received by these funds are included in personal income. Benefit payments made to households by these funds are canceled out as constituting transfers within the household sector and do not show up in the national income and product accounts.

contributions made by property to current production. That is, these payments are deducted from total interest payments to arrive at net interest, which is included in national income (see Table 2–2, page 31). However, since interest payments by government and by consumers are received by households, either directly or through intermediary channels, they have to be added to national income in the process of deriving personal income. In other words, like government transfer payments to persons, they are treated as income received by households but not earned.

The reasons for excluding interest paid by government and consumers from national income need to be spelled out. In the first place, not all borrowing by government and consumers is for the acquisition of capital assets that contribute to current production. Governments may borrow to cover current running expenses, such as the payment of the salaries of police and other public servants, and funds borrowed by the federal government have often been used to purchase military equipment which is not productive in any direct sense and much of which, in any case, is not yielding any current flows of services since it was destroyed in the process of fighting past wars. Consumers borrow to buy goods classified as nondurable, such as clothing, and to pay for services, such as travel costs of going on vacations. It is true that much government borrowing is for the purpose of financing the acquisition of durable assets, such as schools and highways, which do yield continuing flows of valuable services; and consumers similarly borrow to purchase durable goods, such as automobiles and household goods, which likewise provide flows of services as they are used.[4] However, unlike investments by business concerns in plant and equipment, these acquisitions of durable assets by government and by consumers are not treated as investment in the national income accounts. Since business concerns sell their output on the market, the profits they make, together with their interest payments, provide a market-determined measure of the return on their investment that can be included in national income. Government and consumers, however, do not sell in the market the services that are produced by such capital assets as schools, highways, automobiles, and household goods, so the market provides no measure of the return on their investments. What ideally should be done is to estimate or *impute* a value to the flow of services from government and consumer capital to be included in national income. It is distinctly a defect in the national income accounts that this is not done. However, there is judged to be insufficient basis for arriving at a market value for the services of government and consumer assets to justify such estimates.[5]

---

[4] Of course, governments sometimes use current tax revenues to pay for durable productive assets, and consumers purchase durable goods out of current income.

[5] It should be pointed out that houses are handled in a different way from other household durable goods. Houses, even though occupied by their owners, are treated as a form of investment and a return from owner-occupied houses is estimated and included in the national income accounts. Indeed, owner-occupied houses are treated as business concerns and are included in the business sector of the economy. Interest payments on

It would be possible to use government and consumer interest payments as a measure of the return to government and consumer capital. For reasons indicated above, however, this would not be satisfactory, since changes in such interest payments would not necessarily reflect changes in the productive contributions of capital. If business interest costs rise sharply in a way that does not reflect an increase in the contribution of business capital, profits will decline since interest is deducted as an expense in computing profits. In the case of government and consumers, however, in the absence of an independent measure or estimate of the return on capital, there is no similar protection against the effects of erratic fluctuations in interest payments. Although there is some controversy concerning the treatment of interest payments by government and consumers, in view of the difficulties involved it is the prevailing judgment that they should be excluded from national income, as is the current practice of the Department of Commerce.[6]

Business transfer payments, as we have seen, are payments (including charitable contributions, bad debt losses, etc.) by the business sector to the household sector for which no current productive services are rendered. Hence, they are not included in the national income. However, they are receipts of the household sector and therefore are included in personal income.[7]

Personal income may be obtained by adjusting national income by means of the deductions and additions just discussed. It may also be obtained directly by adding together the various kinds of income received by households. For 1968, the calculation is as follows (in billions):

Wage and salary disbursements	$465.0
Other labor income	24.2
Proprietors' income	63.8
Rental income of persons	21.2
Dividends	23.1
Personal interest income	54.1
Transfer payments	59.2
Less: Personal contributions for social insurance	22.6
Personal Income	$687.9

mortgage debt incurred by households for the purchase of houses are therefore treated as business interest payments and included in the national income. See the discussion on pp. 57–59.

[6] Prior to the revisions in the national income accounting system reported in the Department of Commerce *Survey of Current Business,* August 1965, interest payments by consumers—but *not* by government—were included in national income. The abandonment of this practice caused substantial revisions in national income estimates for earlier years. For this as well as other reasons, national income and product estimates published prior to August 1965 are not consistent with estimates published after that date. The student is therefore cautioned not to use the earlier data.

[7] Business transfer payments are charged as expenses against current output by business firms in computing their profits. Hence, as we have seen, they are included in GNP, deducted from GNP to get national income, and then added back in to get personal income. Government transfer payments, on the other hand, are not included in either GNP or national income.

Wage and salary disbursements plus other labor income, which includes (see Table 2–2, page 31) employer contributions to private pension and welfare funds ($20.1 billion) and other supplements to wages and salaries ($4 billion), represents the portion of compensation of employees that is actually received by the household sector. All of the income of unincorporated businesses (proprietors' income and rental income of persons) is treated as being received by the household sector. This leads to a peculiarity in the estimate of personal saving which is discussed below.[8] Personal interest income includes net interest ($28.0 billion), which is included in national income (see Table 2–2), plus interest paid by government (net) and by consumers ($26.1 billion), which is added in the process of moving from national income to personal income. Transfer payments includes government transfer payments to persons ($55.8 billion) plus business transfer payments ($3.4 billion).

Employer contributions for social insurance ($24.4 billion), which are included in compensation of employees in national income (see Table 2–2), are not included in building up personal income. However, personal (i.e., employee) contributions for social insurance ($22.6 billion) are included as part of the wage and salary disbursements component of personal income. But since employee contributions are usually deducted from employees' paychecks by their employers who then remit them to the government, these amounts are not actually received by employees, and they are therefore deducted in the computation of personal income.

The concept of personal income is rendered a bit fuzzy by the fact that a substantial portion of personal income taxes on wage and salary income is deducted at the source—that is, employers deduct taxes from their employees' wages and remit these taxes directly to the government. Consequently, personal income does not *really* measure the amount of income actually received by households. Moreover, there are further tax liabilities that must be met out of the income after it is received so that consumers are not free to spend all of it. Despite these shortcomings, personal income is one of the most important income measures, chiefly because it is the only major income concept that is capable of being estimated with reasonable accuracy on a *monthly basis.* The monthly estimate of personal income, which constitutes a reasonably satisfactory index of the current flow of purchasing power to consumers, is one of the major indicators of the current state of the economy and is accordingly widely reported and discussed in the newspapers.

Personal income may be either smaller or larger than national income. In the Great Depression of the 1930's, personal income was larger than national income in every year from 1930 through 1939. Since that time, personal income has been smaller than national income each year. In periods of declining business activity, national income tends to drop more than

[8] See pp. 48–49.

personal income, because corporate profits usually drop sharply and social security contributions drop with declining payrolls, while dividends are usually maintained fairly well and government transfer payments for relief and unemployment compensation increase.

DISPOSABLE PERSONAL INCOME

Disposable personal income is the amount of income available to the household sector for spending and saving after taxes have been paid. It is computed by subtracting *personal tax and nontax payments from personal income.* This calculation for 1968 is as follows (amounts in billions):

Personal income	$687.9
Less: Personal tax and nontax payments	97.9
Equals: Disposable Personal Income	$590.0

Personal tax and nontax payments includes taxes paid by individuals (mainly, personal income taxes and taxes on nonbusiness property) together with fees paid to government for specific services. Tax refunds are deducted at the time the refunds are paid.

Disposable personal income is the income concept that is most widely used in studies of household behavior, since it is the amount which households are more or less free to use as they please.[9] That is, it is the amount of income available for spending or saving. The spending of households out of disposable personal income is designated as *personal outlays* and is subdivided under the following three headings: *personal consumption expenditures,* which includes spending on currently produced goods and services and is a major component in building up GNP from the expenditure side; *interest paid by consumers,* which, as explained above, is excluded from GNP and national income; and *personal transfer payments to foreigners.* This last item includes gifts made by households, whether in money (such as remittances of immigrants from Europe to their families back home) or in kind (such as CARE packages distributed to poor families abroad). All such transfer payments are treated as though households gave money to the foreign recipients and the recipients used the money to buy goods and services in the United States. Thus, the purchases are reflected in GNP from the expenditure side through net exports rather than through personal consumption expenditures.[10] The allocation of disposable personal income in 1968 was as follows (amounts in billions):

---

[9] Of course, households have debt payments and other contractual or semicontractual obligations to meet, so that they do not have complete freedom even in the use of "disposable" income.

[10] Prior to the 1965 revisions of the account, personal remittances to foreigners in cash and in kind were included in personal consumption expenditures. For a discussion, see the *Survey of Current Business,* August 1965, pp. 10–12.

Disposable Personal Income		$590.0
Less: Personal outlays		551.6
Personal consumption expenditures	$536.6	
Interest paid by consumers	14.2	
Personal transfer payments to foreigners	0.8	
Equals: Personal Saving		$ 38.4

Personal saving is thus computed as a residual—that is, it is the difference between disposable personal income and personal outlays. Since it is small in magnitude relative to the two quantities whose difference it is, personal saving is estimated somewhat less dependably than many of the other components of national income and product; that is, relatively small errors in either disposable personal income or personal outlays will produce proportionately larger errors in personal saving.[11]

One important peculiarity of personal saving is that it includes the saving of unincorporated business enterprises as well as the saving of households. This is because of the great difficulty of separating satisfactorily the accounts of unincorporated business enterprises from those of their owners. The profits of unincorporated enterprises are included in national income. If the necessary information were available, it would be desirable to deduct the portion of these profits that was retained in businesses from national income in the process of obtaining personal income, thus leaving in personal income only that part of unincorporated business profits that was actually withdrawn from such businesses by their owners. This procedure, if it could be adopted, would parallel the treatment accorded corporate profits, as indicated above. But the information is not available, and consequently no deduction is made for retained profits of unincorporated businesses in adjusting national income to obtain personal income. Thus, as was noted in our earlier discussion, the full amount of profits of such enterprises is included in personal income.[12] Taxes on unincorporated business profits, which are directly subject to the personal income tax, are included in personal tax and nontax payments, and deducted from personal income to get personal disposable income. However, the full amount of aftertax profits of unincorporated enterprises is included in personal disposable income, and consequently when personal outlays are deducted from personal disposable income, the resulting figure, which we call personal saving, contains the saving in the form of retained earnings of unincorporated enterprises. This is important to remember in studying behavioral relationships, since the fac-

---

[11] Suppose, for example, that the true values are disposable personal income $300 billion, personal outlays $280 billion, and personal savings $20 billion. If the estimate of disposable personal income is 1 percent too large, or $303 billion, while the estimate of personal outlays is 1 percent too small, or $277.2 billion, the estimate of personal saving will be $25.8 billion, which is 29 percent too large.

[12] See p. 46 above.

tors determining the saving of unincorporated businesses are probably somewhat different from those determining the saving of individuals.

THE GOVERNMENT SECTOR ACCOUNTS

It will be useful at this point to bring together the various government receipts and expenditures, most of which we have already discussed. The information for 1968 is shown in the accompanying table (amounts in billions):

	Federal Government	*State and Local Governments*	*Total*
Receipts	$176.3	$106.2	$264.1
Personal tax and nontax receipts	$ 79.5	18.4	97.9
Corporate profits tax accruals	38.3	3.0	41.3
Indirect business tax and nontax accruals	18.0	59.9	77.9
Contributions for social insurance	40.5	6.5	47.0
Federal grants-in-aid	....	18.3	(a)
Expenditures	$181.5	$107.6	$270.8
Purchases of goods and services	$ 99.5	$100.7	$200.3
Transfer payments to persons	45.7	10.0	55.8
Transfer payments to foreigners (net)	2.1	....	2.1
Net interest paid	11.6	0.3	11.9
Subsidies less current surplus of government enterprises	4.2	−3.4	0.8
Grants-in-aid to state and local governments	18.3	....	(a)
Surplus or Deficit (−), National Income and Product Accounts	$−5.2	$−1.5	$−6.7

(a) Since "federal grants-in-aid" under receipts of state and local governments is identically equal to "grants-in-aid to state and local governments" under expenditures of the federal government, the amount cancels out and is omitted from both receipts and expenditures in calculating the totals for all levels of government.

As indicated, these accounts show receipts, expenditures, and the surplus or deficit separately for the federal government and for state and local governments and also the totals for all levels of government taken together. The federal government sector account is commonly referred to as the "national income accounts budget" or the "NIA budget." However, there is another concept of the federal government budget which is used in the official budget documents and which differs in certain important respects from the one incorporated in the national income accounts and shown here.[13]

[13] Until recently, there were two concepts of the federal budget in addition to the NIA budget that were in general use: the administrative budget and the consolidated

There are only two items in the government sector accounts that have not already appeared in our discussion. One is *federal grants-in-aid to state and local governments,* which appear as a receipt in the state and local government column and as an expenditure in the federal government column. These are, as the title indicates, grants made by the federal government to state and local governments under a variety of federal programs for the purpose of assisting these governments in providing certain types of public services. Since they are transfers from one level of government to another, they are omitted from both receipts and expenditures in constructing the consolidated account for the entire government sector, as shown in the last column of the above tabulation.

The other item that has not yet been discussed is *transfer payments to foreigners* (*net*), which appears in the federal sector account. This includes government *nonmilitary grants,* both in cash and in kind, calculated on a net basis—that is, taking the difference between U.S. government grants to foreigners and foreign government grants to the United States. To the extent that grants lead to shipments of goods out of the United States, these movements of goods appear as exports, thereby increasing the net exports component of GNP from the expenditure side. Thus, all grants made in kind are entered in exports, and cash grants are reflected in exports to the extent that the money is used to purchase goods in the United States. Military grants are treated in a different way: they are recorded as federal government purchases and are reflected in GNP from the expenditure side through this channel—that is, it is the government purchase of the military equipment rather than its subsequent transfer to foreign nations that is recorded in the accounts.

---

cash budget. In 1967, President Johnson appointed a commission to study the formulation and presentation of the federal budget and to make recommendations. In its report, issued in October 1967, the commission recommended the adoption of a so-called unified budget, differing in some respects from all three of the budgets (administrative budget, consolidated cash budget, and NIA budget) then in use. In the interest of avoiding the confusion that had been created by the simultaneous existence of three different concepts of the federal budget, the commission recommended that henceforth the term "budget" be used solely to refer to the new concept it was recommending. [See *Report of the President's Commission on Budget Concepts* (Washington: U.S. Government Printing Office, 1967).] In presenting the budget for the fiscal year 1969 (covering the period July 1, 1968, to June 30, 1969), the Bureau of the Budget adopted the new budget concept recommended by the commission and sharply deemphasized the use of the older concepts. The unified budget differs from the federal sector of the national income accounts mainly in showing, under a separate heading, receipts and expenditures under various federal lending programs, but there are also a number of other significant differences. For purposes of analyzing the impact of the federal government on economic activity, it is still necessary to use the federal sector account of the national income and product accounting system, and in spite of the strictures of the commission, this account will probably continue to be referred to as the "NIA budget." On the differences between the unified budget and the federal sector account, see *Budget of the United States Government, Fiscal Year 1969* (Washington: U.S. Government Printing Office, 1968), Special Analysis B, pp. 473–82.

FOREIGN TRANSACTIONS

We need to discuss further certain aspects of the handling of foreign transactions in the national income accounts. We have seen that one of the components of GNP from the expenditure side is net exports, which is the difference between exports and imports. We now need to develop a slightly different concept, *net foreign investment,* which is equal to net exports minus personal and government transfer payments to foreigners. For 1968, the relevant information is as follows (amounts in billions):

Exports of goods and services		$50.6
Less: Imports of goods and services		48.1
Equals: Net exports		$ 2.5
Less: Transfer payments to foreigners		2.9
Personal	$0.8	
Government	2.1	
Equals: Net Foreign Investment		$−0.3

Net foreign investment is equal to the net increase in claims held by the United States against foreign countries. This can be seen by observing that the excess of our exports over our imports must be paid for in some form by other countries. Payments may be made by foreign shipments of gold to us, through an increase in our deposits in foreign banks (if payment is made in foreign currencies), through a reduction in foreign deposits in our banks (if payment is made in dollars), by our lending to foreigners through purchase of their securities or other debt claims, or by our acquisition of equity claims (stocks) in foreign business enterprises. Since the United States gold stock is treated as a claim against foreign countries (because we can use it at any time to buy goods or services abroad), all of these operations serve, on balance, to increase our claims against foreign countries in one form or another. However, to the extent that our government makes grants to foreign countries in the form of money, this increases their claims against us, and this amount must be offset against the excess of our exports over our imports to arrive at the net increase in our foreign claims.

The student should note that net foreign investment may be *real* investment in very much the same way as investment in domestic plant and equipment for industry, in the sense that it may increase our future potential income. If we have an export surplus and accumulate foreign claims, the profits from direct investment in foreign plants, the interest on foreign bonds and loans, and the dividends on investment in foreign stocks will add to our future income as a nation and, by permitting us to import more goods and services, may contribute to a higher standard of living in the future just as investment of resources in the construction of plants and equipment for domestic industry would enable us to raise our living standards by producing more at home. Moreover, the accumulation of foreign capital may require the use of domestic resources, just as would the construction of plants for domestic industry. One way to accumulate more foreign claims is to devote more resources to the production of goods for export; another is to produce more goods at home to take the place of imported goods, thus permitting imports to be reduced. Assuming resources are fully employed to begin with, either of these alternatives requires that we reduce the amount of resources devoted to some other purpose—say, current consumption—just as would the allocation of more resources to domestic investment in plant and equipment.

SAVING AND INVESTMENT

It will be useful in connection with our later work to collect together at this point the various elements of saving and investment. The table showing this information for 1968 is as follows (amounts in billions):

Gross saving	$125.9
Personal saving	$ 38.4
Undistributed corporate profits	26.7
Capital consumption allowances	73.3
Government surplus or deficit (−), national income and product accounts	−6.7
Corporate inventory valuation adjustment	−3.2
Statistical discrepancy	−2.5
Gross investment	$125.9
Gross private domestic investment	$126.3
Net foreign investment	−0.3

Saving represents the portion of income not used for consumption. Personal saving includes net saving by households and unincorporated businesses. Undistributed corporate profits is net saving by corporations. Capital consumption allowances are amounts set aside out of income to cover destruction and depreciation of existing capital assets. It is important to notice that *gross saving and gross investment are exactly equal.* This equality is easy—although somewhat cumbersome—to explain. From the fact that GNP is equal to the sum of expenditures by the four sectors, we can easily derive the following equation:

$$\text{Gross national product} - \text{Personal consumption expenditures} - \text{Government purchases of goods} = \text{Gross private domestic investment} + \text{Net exports of goods and services} \quad (1)$$

Summarizing the various additions and subtractions that must be made from GNP to get, successively, national income, personal income, personal disposable income, and personal saving, we have the following (see Table 3–1 where all of these adjustments are collected together):

Gross national product − Capital consumption allowances − Indirect business tax and nontax liabilities − Business transfer payments − Statistical discrepancy + Subsidies less current surplus of government enterprises − Corporate profits tax liability − Undistributed corporate profits − Corporate inventory valuation adjustment − Contributions for social insurance + Government transfer payments to persons + Net interest paid by government + Interest paid by consumers + Business transfer payments − Personal tax and nontax payments − Personal consumption expenditures − Interest paid by consumers − Personal transfer payments to foreigners = Personal saving (2)

Rearranging the terms of this equation (omitting *business transfer payments* and *interest paid by consumers,* each of which appears once with a minus sign and once with a plus sign) and subtracting *government purchases of goods and services* from both sides, we have

Gross national product − Personal consumption expenditures − Government purchases of goods and services = Capital consumption allowances + Indirect business tax and nontax liabilities + Statistical discrepancy − Subsidies less current surplus of government enterprises + Corporate profits tax liability + Undistributed corporate profits + Corporate inventory valuation adjustment + Contributions for social insurance − Government transfer payments to persons − Net interest paid by government + Personal tax and nontax payments + Personal transfer payments to foreigners + Personal saving − Government purchases of goods and services (3)

We can now set the right-hand side of Equation 1 equal to the right-hand side of Equation 3, since the left-hand sides to the two equations are the same. This gives us:

Gross private domestic investment + Net exports of goods and services = Capital consumption allowances + Indirect business tax and nontax liabilities + Statistical discrepancy − Subsidies less current surplus of government enterprises + Corporate profits tax liability + Undistributed corporate profits + Corporate inventory valuation adjustment + Contributions for social insurance − Government transfer payments to persons − Net interest paid by government + Personal tax and nontax payments + Personal transfer payments to foreigners + Personal saving − Government purchases of goods and services (4)

Subtracting government transfer payments to foreigners from both sides and collecting together the government receipt and expenditure items on the right-hand side, we obtain the following:

> Gross private domestic investment + (Net exports − Government transfer payments to foreigners − Personal transfer payments to foreigners) = Personal saving + Undistributed corporate profits + Capital consumption allowances + Corporate inventory valuation adjustment + Statistical discrepancy + (Indirect business tax and nontax liability + Corporate profits tax liability + Contributions for social insurance + Personal tax and nontax payments − Government purchases of goods and services − Subsidies less current surplus of government enterprises − Government transfer payments to persons − Government transfer payments to foreigners − Net interest paid by government) (5)

Since the items enclosed in parentheses on the left-hand side are equal to net foreign investment and those enclosed in parentheses on the right-hand side are equal to government surplus or deficit, income and product account, this expression simplifies to

> Gross private domestic investment + Net foreign investment = Personal saving + Undistributed corporate profits + Capital consumption allowances + Government surplus or deficit, national income and product accounts + Corporate inventory valuation adjustment + Statistical discrepancy (6)

This is the relationship shown in the above table of gross saving and investment. Since the corporate inventory valuation adjustment is merely a technical correction and the statistical discrepancy represents errors in data and estimating procedures, the essence of this equation is that gross saving (personal saving, gross business saving, and net saving through the government budget) equals gross investment (gross investment by domestic business concerns and net investment in foreign claims).

It can be seen from the above derivation that gross saving equals gross investment simply because each of them is equal to unconsumed output—i.e., the difference between gross national product and the sum of personal consumption expenditures and government purchases of goods and services. This is the *ex post identity between saving and investment* to which we shall again turn our attention later in this book. As we shall see, however, there is another sense in which saving may differ from investment; in fact such differences will play an important role in our later analysis.[14]

Table 3–1 shows the national income and product accounts, organized along the lines of our discussion, for selected years from 1929 to 1968.

[14] See Chapter 5.

TABLE 3-1

National Income and Product Accounts, Selected Years
(amounts in billions)

	1929	*1933*	*1940*	*1947*	*1961*	*1966*	*1967*	*1968*
1. Personal consumption expenditures	$ 77.2	$45.8	$70.8	$160.7	$335.2	$466.3	$492.3	$536.6
2. Gross private domestic investment	16.2	1.4	13.1	34.0	71.7	121.4	116.0	126.3
3. Net exports of goods and services	1.1	0.4	1.7	11.5	5.6	5.3	5.2	2.5
4. Government purchases of goods and services	8.5	8.0	14.0	25.1	107.6	156.8	180.1	200.3
5. Gross National Product	103.1	55.6	99.7	231.3	520.1	749.9	793.5	865.7
6. Less: Capital consumption allowances	7.9	7.0	7.5	12.3	45.2	63.9	68.6	73.3
7. Equals: Net national product	95.2	48.6	92.2	219.1	474.9	685.9	725.0	792.4
8. Less: Indirect business tax and nontax liability	7.0	7.1	10.0	18.4	47.7	65.7	70.1	77.9
9. Business transfer payments	0.6	0.7	0.4	0.6	2.0	3.0	3.2	3.4
10. Statistical discrepancy	0.7	0.6	1.0	0.9	−0.7	−1.0	−1.0	−2.5
11. Plus: Subsidies less current surplus of government enterprises	−0.1	*	0.4	−0.2	1.4	2.3	1.4	0.8
12. Equals: National Income	86.8	40.3	81.1	199.0	427.3	620.6	654.0	714.4
13. Less: Corporate profits tax liability	1.4	0.5	2.8	11.3	23.1	34.3	33.0	41.3
14. Undistributed corporate profits	2.8	−1.6	3.2	13.9	13.5	29.1	25.9	26.7
15. Corporate inventory valuation adjustment	0.5	−2.1	−0.2	−5.9	−0.1	−1.8	−1.1	−3.2
16. Contributions for social insurance	0.2	0.3	2.3	5.7	21.4	38.0	42.4	47.0
17. Wage accruals less disbursements	0	0	0	*	0	0	0	0
18. Plus: Government transfer payments to persons	0.9	1.5	2.7	11.1	30.4	41.1	48.8	55.8
19. Interest paid by government (net) and by consumers	2.5	1.6	2.1	5.5	15.0	22.2	23.6	26.1
20. Business transfer payments	0.6	0.7	0.4	0.6	2.0	3.0	3.2	3.4
21. Equals: Personal Income	85.9	47.0	78.3	191.3	416.8	587.2	629.4	687.9
22. Less: Personal tax and nontax payments	2.6	1.5	2.6	21.4	52.4	75.4	82.9	97.9
23. Equals: Disposable Personal Income	83.3	45.5	75.7	169.8	364.4	511.9	546.5	590.0
24. Less: Personal outlays	79.1	46.5	71.8	162.5	343.3	479.3	506.2	551.6
25. Personal consumption expenditures	(77.2)	(45.8)	(70.8)	(160.7)	(335.2)	(466.3)	(492.3)	(536.6)
26. Interest paid by consumers	(1.5)	(0.5)	(0.8)	(1.1)	(7.6)	(12.4)	(13.1)	(14.2)
27. Personal transfer payments to foreigners	(0.3)	(0.2)	(0.2)	(0.7)	(0.5)	(0.6)	(0.8)	(0.8)
28. Equals: Personal Saving	$ 4.2	$−0.9	$ 3.8	$ 7.3	$ 21.2	$ 32.5	$ 40.4	$ 38.4

NOTE: Items may not add to totals due to rounding.
* Less than $50 million.
SOURCE: *Survey of Current Business*, July 1969.

SOME SPECIAL PROBLEMS

A number of problems which require special treatment arise in the process of the construction of a satisfactory system of national income and product accounts. Some of these were dealt with in the course of our earlier discussion, but several significant problem areas remain to be discussed.

Inventory Valuation Adjustment

In general, capital gains and losses are excluded from the national income and product accounts. These are gains or losses that arise purely out of changes in the market value of existing property. For example, a person who holds shares of General Motors stock or residential property will benefit if these assets rise in value and will be hurt if they decline in value. Such a gain or loss is said to be "realized" if the owner actually sells the asset on the market at a higher or lower price than that at which he bought it. Gains or losses of this kind are omitted from the national accounts whether they are realized or not. The reason for this is that the purpose of the accounts is to measure production, and it is clear that incomes that take the form of capital gains or losses do not reflect any addition to current output.

There is one place where elements that are essentially capital gains or losses tend to get into the accounts and have to be removed. This is in connection with the valuation of changes in inventories. Consider a situation in which prices rise substantially during the year. At the beginning of the year, businessmen have in stock materials which were purchased at the relatively low prices prevailing in the previous year. As they process these materials into finished products during the year, they use up the materials and charge them to expenses at the low prices. When they sell their final products at the higher prices that develop as the year progresses, their profits appear to be high due to the low-appearing costs and the high prices. However, when they restock their supplies of materials, they have to pay the higher prices for these materials. Their profits for the year are overstated, because they are harvesting what are, in effect, capital gains resulting from the rise in prices between the time they bought the materials and the time they sold the finished goods. The increase in inventory stocks during the year is also overstated or the decrease understated. In a year in which prices are falling, exactly the opposite tendencies appear: that is, profits are understated, and the increase in inventories is understated or the decrease overstated, as the case may be. If the price level undergoes successive rises and falls of approximately equal magnitude, the main effect of this phenomenon is to redistribute profits over time, making them appear too high in periods of rising prices and too low in periods of falling prices. However, if the price level rises or falls more or less steadily, profits are consistently overstated or understated.

In the national income accounts, an effort is made to eliminate fictitious inventory profits or losses by means of an *inventory valuation adjustment.* The Department of Commerce estimates the physical change in inventory stocks and values this change at the average prices that prevailed during the year. The inventory valuation adjustment is obtained by subtracting the estimated change in the value of inventories as shown on the books of business from the estimated change in value obtained by multiplying the physical change in inventory stocks by the average price for the year.[15]

The inventory valuation adjustment is added to the change in business inventories as estimated from business records, and it is also added to business profits. Thus, the value of the change in inventories, which enters into the gross national product from the expenditure side, is adjusted and so is the estimate of business profits which enters into national income from the income side. This double correction preserves the consistency between the estimates of gross national product and national income. When prices rise, the inventory valuation adjustment will be negative (due to the way it is computed) so that the addition of it will reduce profits and the change in business inventories.

The inventory valuation adjustment is computed separately for corporations and unincorporated businesses. In the calculation of gross saving and investment (see table on page 52), it is necessary to add the *corporate inventory valuation adjustment* explicitly to the saving side. This is because it is not reflected in corporate retained earnings (which are as estimated directly from corporate records) while it has already been incorporated in the estimate of the inventory component of gross private domestic investment. Thus, it is necessary to add the corporate inventory valuation adjustment on the saving side in order to maintain the equality between gross saving and gross investment. The inventory valuation adjustment for unincorporated enterprises, on the other hand, is reflected in personal saving, so no specific adjustment is necessary for this.[16]

Imputations

For the most part, the transactions included in the national income accounts reflect purchases and sales of goods and services for money, or other monetary transactions. However, some transactions in kind are in-

[15] To take a simple example, suppose a business held an inventory of a certain product of 100 units valued at $1 per unit for a total book value of $100 at the beginning of the year, while at the end of the year it held 150 units valued at $2 a unit for a value of $300. Thus, the increase in book value during the year would be $200. Suppose, however, that the average price during the year was $1.50. The increase in physical inventory was 50 units, which at a unit value of $1.50, amounts to $75. In this case, the inventory valuation adjustment would be $—125 ($75 — $200 = $—125).

[16] The inventory valuation adjustment involves some complexities not discussed here due to the fact that there are various methods used by business to value inventories. See the Department of Commerce, *National Income: A Supplement to the Survey of Current Business,* 1954 edition, pp. 44–45.

cluded, such as payments for labor services—as when a farmer provides board and lodging to a farm worker. The estimated monetary value of farm produce consumed on farms is included.

Some rather complex procedures are employed in handling residential housing. As we have seen, expenditures on new housing are included in gross private domestic investment. Depreciation of the existing stock of houses is included in capital consumption allowances. To the extent that houses are rented by their occupants from their owners, the payments of rent are included in personal consumption expenditures, and the net rental income after expenses is included in national income. In the case of owner-occupied houses, the rental is estimated or *imputed* on the basis of market rental information for similar houses. The gross imputed rent of owner-occupied houses is included in personal consumption expenditures (and thereby in the gross national product), and the net imputed rental is included in national income, personal income, and disposable personal income. Depreciation on owner-occupied houses is included in capital consumption allowances. In other words, the treatment is as though the owner-occupants paid themselves rent—i.e., they are treated like landlords and included in the business sector of the economy. Accordingly, interest paid on home mortgages is treated as business interest payments and is therefore included in national income—in contrast to interest paid by consumers, which, as explained earlier, is excluded.[17]

Consumer durable goods other than houses—that is, automobiles, furniture, electric appliances, etc.—are treated differently. Purchases of these items are treated as consumption, whereas purchases of houses are treated as investment and as a use of consumer saving. When automobiles are rented through drive-it-yourself agencies, the payments made are treated as personal consumption expenditures in the same way as rental payments for houses. But no effort is made to impute a rental value to owner-driven automobiles. One reason for the difference between the treatment of houses and the treatment of other durable goods is the fact that the practice of renting is much more common in the case of housing than in the case of other durable goods. For one thing, this means that there is a wealth of rental information available on which to base estimates of the rental value of owner-occupied houses. For other durable goods, the rental market is so small that there is frequently insufficient evidence on which to base reasonable estimates of rental values. More important, however, is the fact that when renting is as common as it is in the case of houses, the estimates of national income and product could be distorted significantly by shifts between rental and owner-occupancy. In the absence of the inclusion of rental value of owner-occupied housing, if two people decided to rent each other's houses, the national income and product would rise. If there was a trend—as

[17] See the discussion on pp. 43–45 and in footnote 5.

there has been—toward more and more owner occupancy, the growth in national income and product would be understated. In the case of other durable goods, the rental market is so small that distortions of this sort are of little significance.

Nevertheless, it would be desirable to treat other consumer durable goods as capital goods, include them in investment, and impute a value to the flow of services they render to their owners, provided a satisfactory method could be found for doing this. In principle at least, this method of treatment would permit us to judge better the effects on human welfare of changes in economic conditions. For example, in a severe depression, such as occurred in the 1930's, the drop in national income overstates the decline in welfare, because it reflects the sharp drop in the production and purchase of consumer durable goods but it fails to make allowance for the fact that the existing stock of consumer durables continues to supply a flow of services to its owners.

Another area where imputation would be desirable but where it is impracticable due to the absence of a satisfactory basis for estimation is the case of the services rendered by housewives in the home. Services of paid domestic servants are included in the national product, which has led to the frequent observation that when a man marries his housekeeper the national income declines. Inability to deal effectively with this problem may introduce distortions into the estimates, since, for example, to the extent that there is an increasing tendency to eat meals out, have laundry done commercially outside the home, and so on, the growth of the national income and product may be significantly overstated.

Numerous imputations of interest and other payments that do not actually occur are necessary in the treatment of financial institutions. For example, proper handling of commercial banks requires the imputation of interest payments by the banks to holders of checking deposits and of additional payments of service charges to the banks by these depositors.[18]

Handling of Government Activity

The valuation of the services of government in the accounts involves some serious difficulties and the present method of handling it leaves something to be desired. For one thing, as we have already noted, no allowance is included for the value of the services of government-owned capital. Thus, the government's contribution to the national product (its "value added") is simply the value of the services of government employees. Moreover, it should be noted that these services are valued at what the government pays for them, whereas the other components of the national

[18] The problems involved in dealing with commercial banks and other financial institutions are too complex to be treated extensively here. See *National Income: A Supplement to the Survey of Current Business, op. cit.*, pp. 46–48.

product are valued at what the producing unit sells them for on the market. Such a method of valuation is, of course, impossible in the case of government since it does not offer its product for sale. In any case, however, the same market test is not applied to government services that is applied to other goods and services in the economy, and the valuations placed on government services and other products are accordingly not entirely comparable.

Another difficulty is that all government services are treated as final products, whereas in fact some of these services are rendered to the business sector and thus aid business in producing its output. To the extent that this is the case, the national product is overstated because, in effect, the value of the government service is double-counted—once as a product in itself and again as a component of the business output that it helps to produce. To deal with this problem, it would be necessary to impute to the business sector some charge for the value of the services it receives from government. That is, in addition to deducting purchases of materials from other firms in computing the gross product or value added of a business firm, we should make an additional deduction for "purchases" of services from the government. This would reduce the business sector's product and correct the double-counting. However, no feasible or defensible way of estimating or imputing such a charge has been found. In general, there is little relation between the taxes paid by a business and the value of the services rendered by government to that business—with the possible exception of areas, such as highway construction, where the benefit principle of taxation is employed. Thus, in general, it is not appropriate to treat taxes paid by business as payments for government services and deduct them from sales receipts in computing the gross product of the firm.

QUARTERLY ESTIMATES AND SEASONAL ADJUSTMENTS

In addition to annual estimates of national income and product, the Department of Commerce prepares estimates for each quarter of the year. These *quarterly estimates* are very useful, and they have some features that deserve brief discussion.

In their most useful form the quarterly estimates are presented as *seasonally adjusted quarterly totals at annual rates.* To see what this concept means, let us consider an example. For the fourth quarter of 1962, the unadjusted GNP was estimated to be $151.5 billion. This is, of course, the estimated total gross output during that quarter. To convert it to an annual rate, we merely multiply by four, which gives us $606 billion. This is desirable because it permits us to compare annual and quarterly information —that is, we can, if we wish, compare the rate of production in the fourth quarter of 1962 with, let us say, the rate for the entire year 1962. However, there is another problem which arises from the existence of a rather strong seasonal pattern of variation in GNP. For example, there is a normal tendency for GNP to rise noticeably in the fourth quarter of the year as business activity builds up to a crescendo around Christmas time. On the other hand, activity tends to fall off in the first quarter of the year due partly to the rigors of winter weather. Unless some adjustment was made in the estimates of GNP, it would therefore be difficult to make meaningful comparisons between the fourth quarter of a year and the first quarter of the next year. If we found that GNP declined from the fourth to the first quarter, we would not know how to interpret this fact—that is, we would not know whether it might indicate, let us say, the beginning of a serious decline in business activity or whether it merely reflected the usual seasonal ebb that can be expected each year in the first quarter.

In order to deal with this problem, the Department of Commerce constructs indices of seasonal variation and uses them to remove this variation from the quarterly statistics. For example, on the basis of behavior in past years, the Department estimated that as a result of normal seasonal influences, the GNP in the fourth quarter of 1962 could be expected to be 5.94 percent above the average for the year, or that the *seasonal index* for that quarter was 1.0594. To remove the seasonal variation, the estimated GNP is divided by this seasonal index. When this is done, the estimated GNP after adjustment for seasonal variation, expressed as an annual rate, is $572 billion (606 ÷ 1.0594). This estimate can now be compared with reasonable confidence with the annual estimates for 1962 or other years and with other quarters such as the first quarter of 1963.[19]

Quarterly estimates expressed as seasonally adjusted annual rates are available for GNP, national income, personal income, personal disposable income, and the various breakdowns explained in this chapter. In addition, monthly estimates of personal income and its major components (wages and salaries, dividends, transfer payments, and so on) are available on a seasonally adjusted annual rate basis.

USES OF VARIOUS INCOME CONCEPTS

The national income and product statistics are used for a variety of purposes in economics. Which concept is the proper one to employ depends upon the nature of the problem one is studying. In our study in this book, we shall find the breakdown of the GNP as the sum of expenditures on consumption, gross private domestic investment, net exports, and govern-

[19] Before adjustment for seasonal variation, the GNP *declined* from an annual rate of $606 billion in the fourth quarter of 1962 to a rate of $551.2 billion in the first quarter of 1963. However, after seasonal adjustment, GNP shows an *increase* from $572 billion in the fourth quarter to $577.4 billion in the first. The seasonal indices were 1.0594 for the fourth quarter of 1962 and 0.9546 for the first quarter of 1963. The method used to calculate the seasonal indices allows for changes in the seasonal pattern from year to year, so that the indices for the four quarters differ from year to year.

ment purchases of goods and services particularly useful. The reason is that this is a useful way to organize the analysis of total demand in the economy. Moreover, gross rather than net expenditures provide the better measure of the pressure of demand on available resources in the short run. That is, demand for capital goods for replacement as well as expansion may create employment or, if economic capacity is being fully utilized, may tend to drive up prices.

Other concepts are useful for other purposes, however. In our study of consumer behavior, we shall make some use of disposable personal income, and when we come to consider long-run growth, we shall make some use of net national product. The reason for this latter choice is that it is net rather than gross investment that generates added economic capacity and contributes to growth. However, as noted earlier, the usefulness of the net national product and corresponding concept of net investment is impaired by the difficulty of distinguishing new investment from replacement and the unsatisfactory nature of the accounting measures of capital consumption.

Not only the choice of concepts but the ideal formulation or content of a particular concept may depend upon the problem being studied. For example, in studying consumer behavior, it is possible that estimates of disposable personal income and personal consumption expenditures would be more suitable if they did not contain some of the imputed items that are now included in them, such as net rental income from owner-occupied houses which is included in disposable personal income and gross imputed rent of owner-occupied houses, which is included in personal consumption expenditures. On the other hand, for the study of other problems, such as welfare comparisons or the analysis of economic growth, it might be desirable if even more imputed elements—such as the imputed flow of services from the existing stock of consumer durable goods—were included in the accounts. The structure of accounts employed by the Department of Commerce represents a compromise which is useful for many purposes but is perhaps hardly ideal for any particular use.

MONEY VALUES VERSUS REAL VALUES

National income and product are valued at current market prices. Thus, variations in the magnitudes of the concepts from year to year or quarter to quarter represent a combination of changes in physical quantities and changes in prices. This is a serious shortcoming of the accounts for many purposes. However, there are methods available for adjusting the estimates —or at least portions of them—to take out the variation attributable to price changes. In fact, the Department of Commerce makes such adjustments and publishes the adjusted figures for portions of the accounts. However, these adjustments involve the use of price index numbers, and since such index numbers are taken up in the following chapter, our consideration of these adjustments will be postponed until then.

SUMMARY OF INCOME AND PRODUCT CONCEPTS

Gross national product is the total market value of output produced, before deductions for destruction and depreciation of fixed capital. It is the sum of expenditures on final output by the four sectors, households, business, rest-of-the-world, and government; thus

GNP = Personal consumption expenditures + Gross private domestic investment + Net exports of goods and services + Government purchases of goods and services

National income is the total of all incomes earned for the performance of productive services. It may be derived as the sum of all incomes earned in various forms; thus

National income = Compensation of employees + Proprietors' income + Rental income of persons + Corporate profits + Net interest

Net national product is total output valued at market prices after deduction of capital consumption allowances covering depreciation and destruction of fixed capital. Its value as an income measure is somewhat impaired by the lack of a satisfactory method of measuring the consumption of capital. When indirect business taxes are deducted from net national product, a second estimate of national income is obtained. Indirect business taxes (mainly excise and business property taxes) are treated as expenses of business not assignable to particular income shares and are assumed to be passed on by business in the form of higher prices. This alternative method of deriving the national income leads to a second definition: national income is total output valued at factor costs—i.e., what was paid to factors for its production—after deduction of capital consumption allowances. The *statistical discrepancy* is the difference between the estimates of national income derived from the two definitions and measures the consistency but not the accuracy of the income and product estimates.

GNP may be viewed alternatively as the summation of the market value of output produced (or value added—i.e., sales plus inventory accumulation minus purchases from other units) by all producing units in the economy. The gross product of the government sector is equal to the wages and salaries paid by the government to its employees; no estimate is included in income and product to cover the value of services of government property.

Personal income is the amount of income actually received by the household sector. It may be obtained from national income by deducting income earned but not received and adding income received but not earned. Major deductions include those portions of corporate profits not paid to stockholders (undistributed corporate profits and corporate profits tax liability) and payroll taxes levied for social security and paid directly to government by employers (contributions for social insurance), while important additions include pension and welfare payments by governmental units (government transfer payments to persons) and interest paid by government (net) and by consumers (since such interest is not included in national income). *Disposable personal income* is income of households after payment of taxes and is obtained by deducting personal tax and nontax payments from personal income. Disposable personal income may be either spent or saved. Personal saving is obtained as a residual by deducting personal outlays from personal disposable income. Personal outlays include personal consumption expenditures, interest paid by consumers, and personal transfer payments to foreigners. Personal saving includes the saving of unincorporated businesses as well as that of households and institutions, such as charitable organizations and private pension and welfare funds, which are included in the household sector for want of a better place to put them.

Net foreign investment is the net increase in holdings of claims against foreign countries or their residents. It is obtained by deducting personal and government transfer payments to foreigners from net exports of goods and services, which is the difference between exports and imports. The *government surplus on income and product account* is the difference between total receipts and total expenditures of government units (federal, state, and local) as shown in the national income and product accounts.

Gross saving is the sum of net saving by households, retained earnings of business, allowances for capital consumption, and government saving through the budget (i.e., the government surplus). *Gross investment* includes gross private domestic investment (expenditures for both replacement and expansion of fixed capital—plants, equipment, etc.—by business, expenditures for the construction of new houses, and the change in business inventories), together with net foreign investment. In the national income accounts, gross saving is exactly equal to gross investment, since each is equal to unconsumed output. Thus, subject to minor adjustments and errors, we have

Personal saving + Undistributed corporate profits + Capital consumption allowances + Government surplus or deficit (−)
= Gross private domestic investment + Net foreign investment

Refinements in the accounts designed to make them more satisfactory measures of income and product include the *inventory valuation adjustment* which is designed to eliminate those portions of business profits and of the change in business inventories which result from changes in prices during the period and thus reflect essentially capital gains or losses; also *imputations* of value where no actual market transactions take place in order to reflect incomes received "in kind" in the accounts. An especially notable example of this latter adjustment is the estimation and inclusion of the imputed rental value of owner-occupied houses. This is consistent with the fact that expenditures on new residential construction are included in investment, whereas all other household expenditures on goods and services are included in consumption. *Special problems* arise, to which no entirely satisfactory answers have been found, in connection with the handling of consumer durable goods other than houses, valuation of services rendered by housewives in the home, and the treatment of government activity. In this last category, it should be noted that the government's contribution to national product includes only the value of services of government employees, with no allowances for the services of government property; that these services are valued at cost rather than at market prices, thus impairing comparability with the values produced in the private sector; and that all government services are treated as final products, whereas some of them are in fact intermediate products which help business to produce its output.

Chapter 4 MEASUREMENT OF EMPLOYMENT, UNEMPLOYMENT, AND THE PRICE LEVEL

The performance of the economy is often judged in terms of its success in achieving high employment and stable prices. Accordingly, measures of employment, unemployment, and the price level are important and widely discussed indicators of economic conditions. In this chapter we shall discuss the problems of constructing such measures and also make reference to certain other statistical data that are used in interpreting economic conditions.

EMPLOYMENT AND UNEMPLOYMENT

Considerable progress has been made in recent years in developing more accurate and comprehensive employment statistics and in speeding up the collection of these vital data so that we may be aware rather quickly of changing labor market conditions.

The Monthly Household Survey

The most comprehensive and widely used information on the current state of employment, unemployment, and the labor force is compiled each month from the Current Population Survey conducted by the Bureau of the Census for the Bureau of Labor Statistics. This information is collected by means of personal interviews with a scientifically chosen sample of households spread throughout the United States. The main results of the survey for December 1968, are shown in Table 4–1.

These estimates are based on interviews with about 50,000 households. Trained interviewers visit each of the households in the sample and obtain the requisite information concerning the employment status of members of the household during a specified week in each month. The results obtained from the sample are then "blown up" to provide estimates covering the entire country through the use of independent estimates of the population.

Employed persons include all those who performed any services during the specified week either as paid employees or in their own business or profession (including farming) or who worked 15 hours or more as unpaid workers in an enterprise operated by a member of the family. Also included are persons who were temporarily absent from their jobs or businesses due to illness, bad weather, vacation, labor-management dispute, or personal reasons. The *unemployed* include all persons who did not work at all during the survey week and (*a*) had engaged in some specific job-seeking activity (going to an employment service, applying to an employer, answering a want ad, etc.) within the past four weeks; or (*b*) were waiting to start a new job within 30 days; or (*c*) were waiting to be recalled from layoff. The

TABLE 4–1

Noninstitutional Population, Labor Force, Employment and Unemployment, December, 1968
(thousands of persons 16 years of age and over)

	Unadjusted	*Seasonally Adjusted*
Noninstitutional population	136,619	
Less: Not in labor force	54,001	
Equals: Total labor force	82,618	82,868
Less: Armed forces	3,500	
Equals: Civilian labor force	79,118	79,368
Less: Employment	76,700	76,765
Agricultural	3,279	3,842
Nonagricultural	73,421	72,923
Equals: Unemployment	2,419	2,603
Unemployment rate (unemployment as percentage of civilian labor force)	3.1%	3.3%

NOTE: Items may not add to totals due to rounding.
SOURCE: Department of Labor, Bureau of Labor Statistics.

civilian labor force includes all those classified as either employed or unemployed on the above definitions. The total labor force also includes members of the armed forces. The group classified as not in the labor force includes all members of the noninstitutional population (i.e., persons other than inmates of penal and mental institutions, homes for the aged, and so on) who are not classified as employed or unemployed—that is, persons engaged in housework in their own homes, and those in school, unable to work, or not seeking employment for any other reason. All of the estimates include only persons 16 years old and over, since relatively little work is done by persons under this age due to legal restrictions. Both employment and the size of the labor force are subject to strong seasonal variations. To eliminate these seasonal influences, thereby facilitating meaningful month-to-month comparisons, the Bureau of Labor Statistics prepares and publishes seasonally adjusted estimates such as those shown in the second column of Table 4–1.[1]

[1] For a brief explanation of the nature of seasonal adjustment, see Chapter 3.

The *unemployment rate*—unemployment as a percentage of the civilian labor force—is perhaps the best single index of the state of employment. This ratio is watched very carefully by observers of business conditions. Indeed, the *seasonally adjusted unemployment rate* is perhaps the single most important index of current economic conditions.

In addition to the broad aggregates discussed above, further detailed information is provided, some of it on a regular monthly basis and some less frequently. The regular monthly data include estimates of employment and unemployment by color, age group, sex, hours of work, duration of unemployment, and the occupational distribution of employment.[2]

The monthly sample interview survey is carried out in 449 sample areas including 863 counties and independent cities covering every state and the District of Columbia. A stratified random area sampling technique is employed which is scientifically designed to represent the civilian noninstitutional population 16 years and over, and which permits the estimation of sampling errors.[3]

Table 4–2 presents estimates of the civilian labor force, employment, unemployment, and the unemployment rate as defined in the household survey, for each year from 1929 to 1968. The wide variation in the unemployment rate is quite striking. In the Great Depression of the 1930's, unemployment exceeded 20 percent of the labor force for several years and averaged approximately 18 percent for the entire decade 1930–39. During the second World War, unemployment reached extremely low levels, amounting to only 1.2 percent of the labor force in 1944. Since the war, the rate has averaged about 4.7 percent, but there have been significant increases in the recession years 1949, 1954, 1958, and 1961. There is, of course,

---

[2] Detailed results of the surveys are published each month in *Employment and Earnings and Monthly Report on the Labor Force,* U.S. Department of Labor, Bureau of Labor Statistics. The basic information on employment and unemployment is widely publicized in the newspapers and appears each month in the *Federal Reserve Bulletin,* the *Survey of Current Business,* and other publications.

[3] In November 1961, President Kennedy appointed a Committee to Appraise Employment and Unemployment Statistics under the chairmanship of R. A. Gordon. In September 1962, the committee completed its work and presented its report entitled, *Measuring Employment and Unemployment* (Washington: U.S. Government Printing Office, 1962). This document, often referred to as the Gordon Report, is a comprehensive survey and appraisal of employment and unemployment statistics. While expressing a generally highly favorable opinion of the monthly survey of employment and unemployment described above, the Gordon Report made a number of suggestions for improvement. On the basis of these recommendations, a number of changes were instituted beginning with the January 1967 survey. These changes are described in *Employment and Earnings and Monthly Report on the Labor Force,* February 1967, pp. 3–13. The major changes included an increase in the age limit used in the survey from 14 to 16 years, an increase in the size of the sample from 35,000 households to about 50,000, and certain relatively minor changes in the definitions of employment and unemployment. A technical note describing the procedures employed in carrying out the survey is included each month in *Employment and Earnings and Monthly Report on the Labor Force.* See, for example, the June 1969 issue, pp. 101–14.

TABLE 4–2

Employment and Unemployment, 1929–68

Year	*Civilian Labor Force*	*Employment*	*Unemployment*	*Unemployment as Percentage of Civilian Labor Force*
	Thousands of persons 14 years of age and over			
1929	49,180	47,630	1,550	3.2%
1930	49,820	45,480	4,340	8.7
1931	50,420	42,400	8,020	15.9
1932	51,000	38,940	12,060	23.6
1933	51,590	38,760	12,830	24.9
1934	52,230	40,890	11,340	21.7
1935	52,870	42,260	10,610	20.1
1936	53,440	44,410	9,030	16.9
1937	54,000	46,300	7,700	14.3
1938	54,610	44,220	10,390	19.0
1939	55,230	45,750	9,480	17.2
1940	55,640	47,520	8,120	14.6
1941	55,910	50,350	5,560	9.9
1942	56,410	53,750	2,660	4.7
1943	55,540	54,470	1,070	1.9
1944	54,630	53,960	670	1.2
1945	53,860	52,820	1,040	1.9
1946	57,520	55,250	2,270	3.9
1947	60,168	57,812	2,356	3.9
	Thousands of persons 16 years of age and over			
1947	59,350	57,039	2,311	3.9%
1948	60,621	58,344	2,276	3.8
1949	61,286	57,649	3,637	5.9
1950	62,208	59,920	3,288	5.3
1951	62,017	59,962	2,055	3.3
1952	62,138	60,254	1,833	3.0
1953	63,015	61,181	1,834	2.9
1954	63,643	60,110	3,532	5.5
1955	65,023	62,171	2,852	4.4
1956	66,552	63,802	2,750	4.1
1957	66,929	64,071	2,859	4.3
1958	67,639	63,036	4,602	6.8
1959	68,369	64,630	3,740	5.5
1960	69,628	65,778	3,852	5.5
1961	70,459	65,746	4,714	6.7
1962	70,614	66,702	3,911	5.5
1963	71,833	67,762	4,070	5.7
1964	73,091	69,305	3,786	5.2
1965	74,455	71,088	3,366	4.5
1966	75,770	72,895	2,875	3.8
1967	77,347	74,372	2,975	3.8
1968	78,737	75,920	2,817	3.6

SOURCE: Department of Labor, Bureau of Labor Statistics.

always some unemployment—related, for example, to the tendency of workers to move about and shift from one job to another. The question of what constitutes an acceptable norm for unemployment is one to which we shall devote some attention later in this book.

Other Measures of Employment and Unemployment

The monthly household survey described above has come to be our most reliable tool for assessing the state of employment and unemployment throughout the nation. Other information is available, however, which is also valuable. The Bureau of Labor Statistics collects a great deal of information concerning wage and salary employment in different industries, together with statistics covering average weekly hours of work, average hourly and weekly earnings, and labor turnover, through a monthly sample of voluntary reports by employers. The information derived from this so-called *establishment survey* is not entirely comparable with the data obtained from the household survey referred to above.[4] However, the establishment survey provides a great deal of additional information that is exceedingly useful in interpreting business conditions.

The state unemployment compensation programs provide information concerning unemployment which is less comprehensive than the household survey but which has the advantage of being available on a weekly basis. Estimates are available covering *insured unemployment*—that is, the number of persons covered by the unemployment compensation programs who are unemployed in the particular week. This number is, of course, smaller than total unemployment because some portions of the labor force—including agricultural workers, domestic servants, employees of nonprofit organizations and state and local governments, self-employed persons, some employees of very small firms, new workers who have not earned rights to unemployment insurance, and those who have exhausted their unemployment benefits—are not covered by unemployment compensation. In December 1968, insured unemployment under state unemployment insurance programs averaged 1,172,000 workers, or 48.4 percent of total unemployment as estimated from the household survey. In addition, the unemployment insurance programs provide weekly information concerning new applications for unemployment compensation, which often give early clues to changes in the trend of unemployment.

---

[4] The household survey includes self-employed persons as well as wage and salary workers and covers both farm and nonfarm employment. The establishment survey covers only wage and salary workers on the payrolls of nonfarm establishments. Since the information is derived from employer records, the establishment survey provides no information concerning unemployment. The household survey would count a person merely as employed whether he held one, two, or more jobs. The establishment survey would count him each time he appeared on a payroll.

THE MEASUREMENT OF PRICE LEVEL CHANGES

An *index number* of prices is a device for combining movements of many individual prices for the purpose of estimating the average movement of some specified group of prices.

Different groups of prices move in different ways, and the significance of such movements varies from one group to another. For example, changes in consumer goods prices, which make up the "cost of living" of the population have a different significance from changes in prices of raw materials and semifinished products, which may be reflected in changes in industrial costs in later stages of production. The approach to the measurement of price changes is to use several alternative measures or indexes of such changes. There are various questions that might be asked concerning changes in prices, and an index number should be designed to answer a fairly specific question.

Hypothetical Example

As a starting point for our discussion, suppose we are interested in the behavior of the prices of grain products between 1955 and 1963. We select wheat, corn, rye, and barley as four typical grains and collect the following information concerning their prices and the quantities of them consumed in the two years.[5]

	1955		*1963*	
	Price	*Quantity Consumed (Million Bushels)*	*Price*	*Quantity Consumed (Million Bushels)*
Wheat	$1.75	100	$2.10	90
Corn	1.20	90	1.80	120
Rye	0.50	30	1.50	50
Barley	1.00	20	0.80	40

We want to compute an index number of grain prices for 1963 using 1955 as our *base period.* The simplest way to do this would be to calculate a *price relative* for each of the four grains and average these price relatives together. The price relatives are shown in the following table.

[5] We use hypothetical rather than actual price and quantity data in this illustration in order to simplify the calculations and to bring out more clearly certain aspects of index number construction.

	Price Relative (*1963 Price* ÷ *1955 Price* × *100*)
Wheat	120
Corn	150
Rye	300
Barley	80

The price relatives are really index numbers for each of the individual commodities and merely express the price of each commodity in 1963 as a percentage of its price in 1955—that is, the price of wheat in 1963 was 120 percent of its price in 1955, and so on. We get our index number by simply adding together the four price relatives and dividing by four. Thus the index number is 162.5 (650 ÷ 4). Since the index number for the base year (1955 in this case) is always 100, this calculation suggests that grain prices were 62.5 percent higher in 1963 than in 1955.

The index number just computed is sometimes referred to as an *unweighted* mean of relatives index number, although a more accurate designation would be *equally weighted.* Although this is a perfectly legitimate index number, a glance at the tables presented above raises serious doubt concerning its adequacy as a measure of the change in grain prices. The price relative for rye is much larger than any of the others and when equal weights are employed, as in this index number, this price relative exercises a strong upward pull on the index number. However, the amount of rye consumed is much smaller than the quantities of wheat and corn, which suggests that it should not have an equal weight in the computations.

In order to make allowances for the relative importance of the commodities, we can compute a *weighted aggregative* index number. If we do this using 1955 quantities as weights, the computation is as follows:

$$P_{1963} = \frac{2.10 \times 100 + 1.80 \times 90 + 1.50 \times 30 + 0.80 \times 20}{1.75 \times 100 + 1.20 \times 90 + 0.50 \times 30 + 1.00 \times 20}$$

$$= \frac{433}{318} = 136.2$$

The index number of 136.2 indicates that the average price of grains has risen by 36.2 percent between 1955 and 1963. This is considerably lower than our previous estimate, chiefly because rye is given a reduced weight commensurate with its importance in consumption in 1955. This index number employs *base-year weights* and has an exact meaning which can be stated in words as follows: "It would have cost consumers 36.2 percent more in 1963 than it did in 1955 to have purchased the quantities of these grains that they actually consumed in 1955."

Using the data given above, we can also compute a weighted aggregative index using 1963 quantities as weights. The computation of this index number would be as follows:

$$P_{1963} = \frac{2.10 \times 90 + 1.80 \times 120 + 1.50 \times 50 + 0.80 \times 40}{1.75 \times 90 + 1.20 \times 120 + 0.50 \times 50 + 1.00 \times 40}$$

$$= \frac{512.0}{366.5} = 139.7$$

The meaning of this index number is as follows: "It cost consumers 39.7 percent more in 1963 to purchase the quantities they bought in that year than it would have cost them to buy the same quantities in 1955." This index number employs given-year weights—that is, the quantities used are those consumed in the year for which the index number is being computed. An examination of the data in the tables should convince the student that the reason this index number is larger than the base year weight index (136.2) is that corn and rye, which showed the largest price increases, were consumed in relatively larger quantities in 1963 than in 1955 and therefore receive a heavier weight when 1963 quantities are used than when 1955 quantities are employed as weights.

The two aggregative index numbers just computed occupy an important place in index number theory. The *base-year weight* index is given by the following formula:

$$P_g = \frac{\Sigma p_g q_b}{\Sigma p_b q_b}$$

where the subscript g stands for the "given" year (i.e., the year for which the index number is being computed), the subscript b stands for the base year, p stands for price, q for quantity, and P for index number. The symbol Σ ("sigma") means "take the sum of." Thus, the numerator of the above formula says, "Multiply the given-year price of each commodity by its base-year quantity and sum these products for all commodities," while the denominator says "Multiply the base-year price of each commodity by its base-year quantity and sum these products for all commodities." Similarly, the formula for the *given-year weight* index is as follows:

$$P_g = \frac{\Sigma p_g q_g}{\Sigma p_b q_g}$$

The reader should note that in these index numbers it is the relative rather than absolute quantities that matter. Thus we could divide all the quantities through by 10 (or any other number) without affecting the size of the index number.

In the examples worked out above, the formulas take the following form, since g is 1963 and b is 1955):

$$P_{1963} = \frac{\Sigma p_{1963} q_{1955}}{\Sigma p_{1955} q_{1955}} \qquad \text{Base-year weight index}$$

$$P_{1963} = \frac{\Sigma p_{1963} q_{1963}}{\Sigma p_{1955} q_{1963}} \qquad \text{Given-year weight index}$$

Weighted aggregative index numbers can be computed using weights other than base-year or given-year quantities. For example, an index for 1963 could be computed using 1955 as base year and using 1959 quantities as weights (although, of course, to do this we would need to have the 1959 quantities, which are not given in the above tables). The formula for such an index number would be

$$P_{1963} = \frac{\Sigma p_{1963}q_{1959}}{\Sigma p_{1955}q_{1959}}$$

If we were to compute this index number and it turned out to be, let us say, 137.3, its meaning in words would be: "It would have cost 37.3 percent more in 1963 than in 1955 to have bought the quantities of these commodities that consumers purchased in 1959." There is thus great flexibility in the choice of weights; they can be quantities of base year, given year, some other year, the average of several years, or hypothetical weights based on careful study of consumption habits.

The student should note that, in general, the choice of base period and the choice of weights are independent of each other. The base period appears as the subscript attaching to prices in the denominator, while the weight designation appears as the subscript attaching to quantities in both the numerator and denominator. Thus the general formula for a weighted aggregative index is

$$P_g = \frac{\Sigma p_g q_w}{\Sigma p_b q_w}$$

where g stands for given year, b stands for base year, and w stands for weight designation.

In addition to weighted aggregative index numbers, weighted mean of relatives indexes may be computed, taking the price relatives such as those computed above and weighting them to obtain an average. We shall not discuss this type of index number, however, since as a practical matter most index numbers now in use are of the weighted aggregative type.

Problems of Index Number Construction

It is possible to distinguish several broad problems that are encountered in the process of constructing a price index. We shall take up these problems in the next few paragraphs.

1. ***What Is Being Measured?*** The first and most important decision that must be made concerns the fundamental purpose of the index number. The present view is that there are many index numbers that may be constructed and that an index number should be designed to answer a fairly specific question. Perhaps we want to measure the behavior of the price level of consumer goods as it relates to some fairly specific group of people. Or

maybe we are interested in price formation at earlier stages of production because we feel that these prices become costs of production in later stages and will eventually be reflected in prices charged to consumers. On the other hand, we may be interested in construction costs or prices of producers' durable goods or prices received by farmers. Whatever our objective is, it should be defined as clearly as possible and our index number should be designed to meet it.

2. Choice of Items to Be Included. The kinds of things whose prices are to be included in the index are pretty well determined once the objective of the index has been decided upon. For example, if we want to measure the cost of living as it relates to some class of people, we want to include prices of goods consumed by this group. However, the number of specific items eligible for inclusion may run into the thousands, and it is almost always necessary to use some kind of sampling procedure to select the specific items to be incorporated in the index. Ordinarily a random sampling technique is not employed, but rather a systematic effort is made to include items that are typical of the different classes of goods covered in terms of quantitative importance and also of price behavior.

3. Selection of Weights. As indicated earlier, a very wide choice of possible weights may be employed. One question is whether we should use *fixed weights* or *weights that change continuously.* On the one hand, we might select the quantities consumed or produced in the base year or some other past year and use these each year in computing the index; on the other we might use the quantities of the given year as weights.

There are two strong reasons for using fixed weights, and almost all indexes do so in practice. One reason is the practical one that the use of changing weights involves a great deal more work, since it is necessary to be collecting new quantitative information all the time. In fact, this problem is so serious that indexes with shifting weights are usually wholly impracticable. In addition to this, however, there is a difficulty involved in the interpretation of an index with changing weights. Suppose, for example, that we have an index having 1955 as base year and that this index has a value of 102 for 1960 and a value of 107.1 for 1961. Thus, the value of the index increased by 5 percent between 1960 and 1961 (since 107.1 is 105 percent of 102). If the index has fixed weights, it is possible to give a precise interpretation to this increase. For instance, if 1959 quantities are used as weights, we can say that the index indicates that it cost 5 percent more in 1961 than it did in 1960 to buy the quantities that consumers bought in 1959. On the other hand, if the index employs given-year weights—so that 1960 quantities were used in computing the 1960 index and 1961 quantities were used in computing the 1961 index—it is nearly impossible to give a precise interpretation to the 5 percent increase. Thus, the use of fixed weights facilitates the making of meaningful comparisons between two years neither of which is the base year.

4. Selection of a Base Period. The choice of a base period or base year is of no fundamental importance. This is because price indexes do not tell us whether prices are high or low in any *absolute* sense but permit us to estimate only their *relative* change from one year to another. This point is illustrated in the following table, which shows hypothetical index numbers for 1953, 1958, and 1963, using alternatively 1953 and 1958 as base year.

Year	*Index Number (1953 = 100)*	*Index Number (1958 = 100)*
1953	100	80
1958	125	100
1963	160	128

Starting with the index numbers based on 1953, the 1958- based numbers can be derived by simply dividing through by the index number for 1958 on 1953 as base.[6] Hence all the index numbers based on 1953 are 25 percent larger than the corresponding numbers based on 1958. Either set of numbers indicates that the price level rose by 25 percent between 1953 and 1958 and 28 percent between 1958 and 1963. Thus, changing the base year merely raises or lowers the level of the index numbers for the various years, while leaving their relationship to each other unchanged. Since the level is

[6] Thus the numbers in the second column of the table are obtained by dividing the corresponding numbers in the first column by 1.25. (It should be remembered that when index numbers are used in mathematical manipulation they must first be converted from percentages to decimals). In the case of aggregative indexes having fixed weights, conversion from one base year to another by means of division, as explained above, gives exactly the same results as would be given by constructing the index from the start using the new base year provided the same weights are used. For example, an index number for 1963 using 1953 as base with 1957 weights is given by

$$\frac{\Sigma p_{1963}q_{1957}}{\Sigma p_{1953}q_{1957}} \quad (1)$$

The same index number using 1958 as base is given by

$$\frac{\Sigma p_{1963}q_{1957}}{\Sigma p_{1958}q_{1957}} \quad (2)$$

The index number for 1958 on a 1953 base is

$$\frac{\Sigma p_{1958}q_{1957}}{\Sigma p_{1953}q_{1957}} \quad (3)$$

If we have index numbers (1) and (3) on a 1953 base and want to obtain the index number for 1963 with 1958 as the base, we can divide (1) by (3). This will give us (2), which is precisely the result we could get if we computed the index number for 1963 on a 1958 base directly.

not meaningful anyhow and the whole purpose of index numbers is to bring out relative changes over time, the base year really makes no difference.[7]

Usually in selecting a base year, an effort is made to pick a reasonably "normal" year when prices were neither excessively high nor excessively low. This is because some users may erroneously attach meaning to the absolute values of the index number. Another and more important factor to be taken into account in choosing a base year is the fact that users frequently have occasion to compare movements of different index numbers and that such comparisons are somewhat facilitated—that is, it is not necessary to effect a change of one of the base years—if the same base period is used for both indexes. For this reason, the period 1957–59 is used as the base period for a number of index numbers computed by government agencies.

Errors and Biases in Index Numbers

Probably the most important matters involved in the construction of a satisfactory index number are the decision as to what is being measured and the selection of the specific sample of items to be included. It is also important to get accurate price quotations which are comparable from one time to another and, so far as possible, to include only items which can be specified in such a way that there is reasonable certainty that the same item is being priced in each period. Errors in the data employed and lack of comparability from one period to another can impair the usefulness of an index number in a serious way. In addition to outright errors such as these, there are certain biases that may be present in the index number which may tend to make it exaggerate price movements systematically in one direction or the other. A factor which tends to make the index number systematically overstate price increases and understate price decreases is said to impart an *upward bias* to the index number, while a factor which tends to understate price increases and overstate price decreases is said to give the index number a *downward bias*. There are several common sources of bias that deserve discussion.

1. Type Bias. In practice, nearly all index numbers, including the types discussed above, employ an ordinary or "arithmetic" average—that is, the items included are, in effect, added together and the index number is obtained by dividing by the number of items. An arithmetic average has a systematic upward bias. To illustrate this, let us take a very simple example. Suppose we have two items of equal importance for which we wish to compute a price index. From the base year to the given year the price of one of the products doubles while the price of the other halves. Thus, the price

[7] Due to adjustments for quality changes, the introduction of new products, and periodic changes that must be made in weights, index numbers are not as "neutral" with respect to changes in the base year as the above discussion indicates. Nevertheless, the most important thing to understand about the base year is that it is relatively unimportant.

relatives are 200 for the first product and 50 for the second. If we take the arithmetic averages of these two price relatives, our price index will be 125. Yet it is fairly clear that if the two products are of equal importance, the price level of the two together has not changed. If we use the *geometric average* instead of the arithmetic average, we will get the correct result.[8] To compute the geometric average, we multiply the two price relatives together, which gives us 10,000 (200 × 50). We then take the square root of the product, which gives us 100, indicating that there has been no change in the level of the two prices taken together. It may be noted that whenever there is any variability at all among the items being averaged, the arithmetic average is necessarily larger than the geometric average. Thus, the use of the arithmetic average imparts an upward bias to an index number. This bias is commonly referred to as type bias. It could be eliminated by the use of a geometric average, but the geometric average is so much more cumbersome to use that it is almost never employed.

2. Biases Connected with Weighting. In computing a consumers' price index, what we would be ideally interested in obtaining would be changes over time in the cost of obtaining a *given level of satisfaction.* Let us suppose to begin with that consumers' tastes are constant over the time span we are considering. Then it can be shown that two weighted aggragative indexes, one using base-year weights and the other using given-year weights, set the upper and lower limits, respectively, to the "true" change in the price level. The aggregative index using base-year weights overstates the price increase or understates the price decrease—in other words, it has an upward bias. The reason for this is that it fails to allow for the fact that the consumer can to some extent shield himself from the effects of price changes by making substitutions in his budget. For example, if prices rise, the consumer can be expected to consume relatively less of those products whose prices rose most and relatively more of those products whose prices rose least (or perhaps even declined). Thus, the base-year weight index attaches too much weight to the items whose prices rose most and too little weight to those whose prices rose least and thus overstates the price increase. The given-year weight index, on the other hand, overestimates the consumer's ability to protect himself by substitution and thereby gives too little weight to items whose prices increased most and too much weight to those whose prices increased least and thus understates the rise in prices. The true price change—i.e., the change in the cost of maintaining a constant level of satisfaction—lies somewhere between the base-year weight and the given-year weight indexes, although there is no way of knowing precisely where the true index lies between these limits.

---

[8] A geometric average of *n* items is computed by multiplying all of the items together and taking the *n*th root of their product. In the case of a weighted average, each of the items is multiplied together as many times as the weight attached to it and the root taken is equal to the sum of the weights. Computation of a geometric average is facilitated by the use of logarithms.

Thus, it is often said that a base-year weight index has an upward bias and a given-year weight index a downward bias. Since a *fixed weight* index of any sort is similar to a base-year weight index in the sense that it makes no allowance for substitution, and since most index numbers used in practice are fixed weight indexes for reasons mentioned earlier, it might seem to follow that most index numbers in practical use have an upward bias due to their weighting.

Unfortunately, however, the effects of weighting are not as simple as the above analysis would suggest. It will be recalled that our reasoning was based on the assumption of constant tastes. In practice this assumption can hardly be even approximately correct. With new products constantly being introduced and the effects of education, advertising, and other forces constantly at work, the tastes of individuals undoubtedly change frequently. Moreover, we are considering groups of people, which means that the very concept of tastes is of dubious validity and that to the extent that it does have meaning it must refer to some kind of composite of the tastes of the group. Group tastes, if they can be said to exist at all, are necessarily changing constantly due to the shifting composition of the group as well as to changes in the incomes of the various members, which must be used as weights in combining individual tastes to get those of the group.

Once changes in tastes are allowed for, the very concept of an index number as a measure of changes in the cost of achieving a constant level of satisfaction breaks down, since it makes little sense to compare levels of satisfaction between two periods in which tastes are different. Moreover, the problem of assessing the biases introduced into index numbers by the use of fixed vs. changing weights becomes both more complicated and more uncertain. Whereas, in our earlier analysis, basic demand conditions were treated as constant (fixed by unchanging tastes) and all price changes were due to changes in supply conditions, we are now faced with a situation in which changes may originate on either the supply side or the demand side or both. To the extent that changes are initiated on the supply side, lower prices will be associated with larger quantities, whereas, when changes originate on the demand side, higher prices will usually be associated with larger quantities. Thus, if changes in supply are the predominant influence, our earlier analysis will hold good—that is, a fixed weight index will have an upward bias and a variable weight index a downward bias. However, if changes in demand predominate, just the opposite will hold true. The conclusions are summarized in the following table:

	Fixed-Weight Index	*Variable-Weight Index*
Changes in supply predominate	Upward bias	Downward bias
Changes in demand predominate	Downward bias	Upward bias

Thus, while the weighting does introduce biases, their direction is uncertain and it is very probable that in the same index the weight bias is in one direction at one time and in the other direction at another time.

3. Other Sources of Bias. Changes in the quality of products and the appearance of new products create problems in the construction of price indexes. Even if these problems are handled with the greatest possible care, they may result in some degree of bias. This kind of bias is discussed below in connection with the evaluation of the major index numbers now in use in the United States.

The Consumer and Wholesale Price Indexes

The two most important and most widely used price indexes in the United States are the Consumer Price Index and the Wholesale Price Index, which are prepared by the Bureau of Labor Statistics of the Department of Labor.

1. The Consumer Price Index. The Consumer Price Index is designed to measure changes in the prices of the goods and services bought by urban wage earners and clerical workers. Estimates are available going back to 1931, although, of course, the index has been revised several times since its beginning. The index is available monthly, and employs the base period 1957–59, with 1960–61 weights. Thus, the index number for March 1969 was computed essentially by use of the following formula:

$$P_{\text{Mar. 1969}} = \frac{\Sigma p_{\text{Mar. 1969}} q_{1960-61}}{\Sigma p_{1957-59} q_{1960-61}}$$

The sample used in preparing the index contains about 400 specific goods and services, classified under five main headings: food, housing (including both rental and home ownership), apparel and upkeep, transportation, and health and recreation. The items to be included and the weights employed are determined on the basis of periodic detailed surveys of family expenditure patterns, the latest of which was conducted during the period 1960–61. While the fundamental structure of the index rests on this survey and 1960–61 weights are accordingly employed, adjustments are frequently made to take account of the appearance of new products and changes in the quality of existing products.

Prices are collected from representative stores and other establishments in 56 cities throughout the country. The sample of cities is drawn to represent a universe of about 3,000 urban places in the United States, including Alaska and Hawaii, and includes 39 Standard Metropolitan Statistical Areas and 17 smaller cities. A separate index is published for each of the 23 largest cities in the sample. The national index and the indexes for five major cities (Chicago, Detroit, Los Angeles, New York, Philadelphia) are published monthly, while the indexes for 18 other standard metropolitan areas are published quarterly. Population weights derived from the 1960 census are used to combine the indexes for the individual cities to get the national index. In addition to the overall index, subindexes are published covering the main classifications of goods and services included in the index, as well as a number of subdivisions within these classifications.[9]

2. The Wholesale Price Index. The Wholesale Price Index differs fundamentally from the Consumer Price Index. It excludes entirely prices paid by household consumers and attempts to measure prices charged in primary markets for transactions in large lots in which both buyer and seller are commonly business concerns. As far as possible, the prices collected are those applicable to the first important commercial transaction in the commodity. Price data are collected monthly covering a sample of about 2,300 commodities, classified under the headings of farm products, processed foods and feeds, and industrial commodities. Industrial commodities, in turn, are subdivided into 13 categories. Separate index numbers are computed for each of these classifications and for smaller groupings of commodities within them.[10] The overall Wholesale Price Index is, in effect, a weighted average of these components. Unlike the Consumer Price Index, the Wholesale Price Index is confined to physical commodities and does not include any services at all.

The base period of the index is the period 1957–59. The weights are designed to account for the value of all commodities sold in the domestic market, including imports. The basic weights employed are 1963 shipment values of commodities as reported in industrial censuses, although adjustments are made from time to time to keep the weights up to date. Essentially, the formula used in computing the index for March 1969 was

$$P_{\text{Mar. 1969}} = \frac{\Sigma p_{\text{Mar. 1969}} q_{1963}}{\Sigma p_{1957\text{–}59} q_{1963}}$$

The index is prepared monthly on a nationwide basis, with subindexes covering different categories of commodities in considerable detail. It is available in a continuous series going back to 1890, although many changes have been made in the index since that time.

---

[9] For a more detailed description of the Consumer Price Index as well as an account of its historical development, see *The Consumer Price Index: History and Techniques,* Bulletin No. 1517, Department of Labor, Bureau of Labor Statistics.

[10] The 13 major categories under industrial commodities are: textile products and apparel; hides, skins, leather, and related products; fuels and related products and power; chemicals and allied products; rubber and rubber products; lumber and wood products; pulp, paper, and allied products; metals and metal products; machinery and equipment; furniture and household durables; nonmetallic mineral products; transportation equipment; and miscellaneous products. For a description of the index and extensive material on its composition and weighting, see *Wholesale Prices and Price Indexes,* January 1967 Final and February 1967 Final, U.S. Department of Labor, Bureau of Labor Statistics.

GNP in Constant Dollars and Implicit Price Deflators

Price indexes are often used to "deflate" dollar magnitudes to take out the effect of price changes. The logic of deflation can best be explained by means of a very simple example. Suppose a family bought 500 packages of cigarettes at $0.20 per package in 1961 and 440 packages at $0.25 per package in 1962. That is, the amount spent on cigarettes rose by 10 percent between 1961 and 1962 (from $100 to $110), while the physical quantity bought fell by 12 percent. The price ratio for cigarettes for 1962 with 1961 as base year (p_{1962}/p_{1961}), expressed in the form of a price index, is 125. Dividing the family's actual expenditure for cigarettes in 1962 ($p_{1962}q_{1962}$) by this price index (converted to a decimal), we obtain $88 ($110 ÷ 1.25 = $88), which is the family's purchases in 1962 valued at the 1961 price (symbolically, $p_{1962}q_{1962} \div p_{1962}/p_{1961} = p_{1961}q_{1962}$). Thus, when the purchases of both years are valued at the 1961 price, the amount declines by 12 percent between 1961 and 1962 (from $100 to $88), reflecting only the change in the physical quantity bought. By applying this procedure to the family's cigarette purchases for several years—i.e., by "deflating" each year's purchases with a "price index" for cigarettes with 1961 as the base year—we could obtain a series giving the quantities for each year valued at the 1961 price. This series would, of course, behave in the same way as the series of physical quantities bought, since the entire variation in it would be due to the variation in quantity.

There is little sense in this procedure as applied to a single homogeneous commodity, of course, since changes in the quantity of it can be measured directly (in packages in this case). However, if we wish to measure physical quantities for heterogeneous aggregates of goods and services, we cannot use the simple, direct method since we cannot add together bushels of wheat, tons of steel, pairs of shoes, and so on. One way to get a total for a heterogeneous aggregate, however, is to value each of its components in dollars, which will permit us to add the items together. Moreover, if we value the various components at constant prices—say, the prices that prevailed in a particular year—we have an aggregate whose variation reflects only changes in physical quantities.[11] Such a deflated aggregate is sometimes said to be expressed "in real terms."

This procedure is used to derive estimates of the GNP and its various expenditure components in constant dollars. The method followed is to break down the GNP into small components and then deflate each of these components with an appropriate price index. We might, for example, take

---

[11] We are here combining actual physical quantities using fixed prices (of some specified period) as weights. Thus, the behavior of the resulting aggregate is going to depend to some extent on the weights used—i.e., whether we use prices of one year or another year.

total expenditures on machine tools for 1968 and divide it by an index number of machine tool prices (e.g., a subindex of the Wholesale Price Index), using 1958 as base year. This division gives us an estimate of machine tool production for 1968 valued at 1958 prices. Calculations of this kind are carried out for all the components of GNP and then the deflated estimates are reassembled to give us an estimate of GNP and its expenditure components valued at 1958 prices. The tabulation for 1968 is as follows (billions of 1958 dollars):

Gross national product	707.6
Personal consumption expenditures	452.6
Gross private domestic investment	105.7
Net exports of goods and services	0.9
Government purchases of goods and services	148.4

A similar breakdown of the GNP in 1958 dollars for selected years since 1929 is given in Table 4–3. The Department of Commerce prepares

TABLE 4–3

Gross National Product and Expenditures Valued at Constant Prices, Selected Years (billions of 1958 dollars)

	1929	*1933*	*1940*	*1947*	*1961*	*1966*	*1967*	*1968*
Gross national product	203.6	141.5	227.2	309.9	497.2	658.1	674.6	707.6
Personal consumption expenditures	139.6	112.8	155.7	206.3	322.5	418.1	430.3	452.6
Gross private domestic investment	40.4	5.3	33.0	51.5	69.0	109.3	100.8	105.7
Net exports of goods and services	1.5	0.0	2.1	12.3	5.1	4.2	3.6	0.9
Government purchases of goods and services	22.0	23.3	36.4	39.9	100.5	126.5	140.0	148.4

NOTE: Details may not add to totals due to rounding.
SOURCE: Department of Commerce.

quarterly estimates of the GNP in constant dollars at seasonally adjusted annual rates.

Conceptually, the GNP valued at current prices for 1968 may be represented as

$$\Sigma p_{1968} q_{1968} \tag{1}$$

where the q's and corresponding p's cover all of the goods and services included in GNP. The deflated GNP, obtained as explained above, by dividing the components of the GNP at current prices by appropriate price indexes having 1958 as base year and the reassembling the components, may be represented as

$$\Sigma p_{1958} q_{1968} \tag{2}$$

If we divide Equation 1 by Equation 2, we get a price index for 1968 with 1958 as base year and 1968 quantities as weights, thus

$$P_{1968} = \frac{\Sigma p_{1968}q_{1968}}{\Sigma p_{1958}q_{1968}}$$

This index number is the ***implicit price deflator of the GNP,*** and, as can be seen, it is simply the GNP at current prices divided by the GNP at constant prices. Implicit price deflators can be worked out for separate components of the GNP, such as personal consumption expenditures, gross private domestic investment, and so on. Table 4–4 gives the implicit price deflators for major

TABLE 4–4

Implicit Price Deflators for 1968
(percent; 1958 = 100)

Gross national product	122.3
Personal consumption expenditures	118.6
Durable goods	103.3
Nondurable goods	117.1
Services	127.3
Gross private domestic investment	(*a*)
Fixed investment	120.0
Nonresidential	117.1
Structures	129.3
Producers' durable equipment	111.9
Residential structures	129.7
Change in business inventories	(*a*)
Net exports of goods and services	(*a*)
Exports	110.9
Imports	107.6
Government purchases of goods and services	135.0
Federal	126.2
State and local	145.0

(*a*) Separate implicit price deflators are not computed for total gross private domestic investment, for change in business inventories, and for net exports of goods and services.

subdivisions of GNP for 1968 (index numbers, with 1958 = 100).

The implicit price deflators are useful by-products of the estimation of GNP in constant dollars and are now available on both a quarterly and an annual basis. They are essentially given-year weighted indexes, although, of course, their nature depends indirectly upon the composition of the detailed index numbers which are used to deflate the components of GNP to get the constant dollar estimates, and these indexes in most cases employ weights from earlier years.

The implicit price deflator for the total GNP is the nearest thing that is now available to an estimate of the general price level. The deflators for individual components of GNP are also useful for many purposes. It should be noted that the deflator for personal consumption expenditures differs from the Consumer Price Index in part because the former measures the price level as it affects all consumers, while the latter is intended to cover only wage earners and clerical workers in urban areas.

Comparison of Major Index Numbers

The values of the Consumer Price Index, the Wholesale Price Index, and the implicit price deflator for the total GNP for each year of the postwar period are shown in Table 4–5.[12] On a monthly basis, the Wholesale Price Index is somewhat more variable than the Consumer Price Index and its movements tend to lead the Consumer Price Index. In part, the explanation is that the Wholesale Price Index reflects price adjustments in early stages of production which enter into costs and affect the prices of consumer goods at a later time. It is not uncommon to find that the two indexes are moving in opposite directions at certain times. Movements of the implicit price deflator are intermediate between movements of consumer prices and wholesale prices; this is due to the fact that many of the components of both the Consumer and Wholesale Price Indexes enter indirectly into the implicit price deflator as explained above.

The economic significance of the Consumer Price Index is considerably enhanced by the fact that it provides the basis for wage escalation provisions of many labor-management contracts. Roughly four million workers are employed under such contracts, by which their wages are adjusted on a quarterly or seminnual basis for changes in the purchasing power of the dollar as measured by the Consumer Price Index. In addition, the index is a factor in nearly every wage negotiation through collective bargaining in the economy.

Shortcomings of Major Price Indexes[13]

All of the major indexes employ an arithmetic average and are therefore characterized by upward type bias, as indicated earlier. However, this bias is

[12] The Consumer Price Index and the Wholesale Price Index are both published each month in the *Survey of Current Business* and the *Federal Reserve Bulletin.* The implicit price deflators are published from time to time in the *Survey of Current Business.*

[13] In 1961, at the request of the Bureau of the Budget, a Price Statistics Review Committee appointed by the National Bureau of Economic Research presented a report containing a comprehensive review and evaluation of the price indexes prepared by the U.S. government. See *The Price Statistics of the Federal Government: Review, Appraisal, and Recommendations* in *Government Price Statistics,* Hearings before the Subcommittee on Economic Statistics of the Joint Economic Committee, 87th Congress, 1st Session, Part 1, January 24, 1961 (Washington: U.S. Government Printing Office, 1961), pp. 9–99. This report, often referred to as the Stigler Report (after George J. Stigler, the chairman of the review committee) is the most valuable, authoritative, and dispassionate evaluation available. The review committee also provided a series of 12 staff papers, prepared by experts on various aspects of price indexes, which were printed along with the report (pp. 101–525).

TABLE 4-5

Leading Price Indexes, 1929–68

Year	*Price Indexes (1957–59 = 100)* Consumer	Wholesale	*Implicit Price Deflator—GNP (1958 = 100)*
1929	59.7	52.1	50.6
1930	58.2	47.3	49.3
1931	53.0	39.9	44.8
1932	47.6	35.6	40.2
1933	45.1	36.1	39.3
1934	46.6	41.0	42.2
1935	47.8	43.8	42.6
1936	48.3	44.2	42.7
1937	50.0	47.2	44.5
1938	49.1	43.0	43.9
1939	48.4	42.2	43.2
1940	48.8	43.0	43.9
1941	51.3	47.8	47.2
1942	56.8	54.0	53.0
1943	60.3	56.5	56.8
1944	61.3	56.9	58.2
1945	62.7	57.9	59.7
1946	68.0	66.1	66.7
1947	77.8	81.2	74.6
1948	83.8	87.9	79.6
1949	83.0	83.5	79.1
1950	83.8	86.8	80.2
1951	90.5	96.7	85.6
1952	92.5	94.0	87.5
1953	93.2	92.7	88.3
1954	93.6	92.9	89.6
1955	93.3	93.2	90.9
1956	94.7	96.2	94.0
1957	98.0	99.0	97.5
1958	100.7	100.4	100.0
1959	101.5	100.6	101.6
1960	103.1	100.7	103.3
1961	104.2	100.3	104.6
1962	105.4	100.6	105.8
1963	106.7	100.3	107.2
1964	108.1	100.5	108.8
1965	109.9	102.5	110.9
1966	113.1	105.9	113.9
1967	116.3	106.1	117.6
1968	121.2	108.7	122.3

SOURCE: Departments of Labor and Commerce.

not very important. Biases are also introduced by the weighting employed, but the direction of these biases depends upon whether changes in supply or in demand are the predominant influence, as we have seen. Hence, it is doubtful whether these biases work systematically in one direction, and they are probably not very important.

Many economists believe, however, that all three of the major price indexes contain significant upward biases—that is, that they tend to overstate increases and understate decreases in the price level. Two major reasons are given for this alleged upward bias. First, the indexes are said not to make adequate allowances for improvements in the quality of goods and services. Second, there is said to be a bias connected with the introduction of new products. Those responsible for the construction of the indexes do make substitutions when new products displace old ones. However, it is argued that there is often too much delay in making such substitutions and that this tends to create an upward bias, because new products are especially likely to decline in price during the early stages of their development as rapid increase in their use permits the introduction of mass production techniques which lower unit costs of production.

The possible bias resulting from changes in quality is especially a source of controversy. In constructing the Consumer Price Index, the Bureau of Labor Statistics attempts to adjust for those improvements in quality which cause upward adjustments in prices; however, there may be improvements which do not raise prices and no adjustments are made for these. The pricing of services involves especial difficulties. For one thing, the compensation of the person rendering the services is ordinarily the major determinant of the price, so that improvements in the quality of the service rendered are not adequately allowed for. A good example is medical services. Advances in medical research have made it possible to cope with health problems that could not be handled at all a few years ago, while in other cases the time taken to cure an ailment is much shorter than it used to be. It is almost impossible to make adequate allowance for such factors as this in compiling price indexes. Similar problems arise in connection with other professional and personal services.

The same kinds of difficulties are present in the Wholesale Price Index, although probably in lesser degree. Producers' durable goods, whose prices are included in the index, are constantly being improved. A machine tool of a particular type today may do its job much more rapidly and at a lower cost in terms of labor and materials than did its counterpart of 10 years ago. It is almost impossible to make proper allowance for such advances.

The implicit price deflator includes most of the biases that are present in the Consumer and Wholesale Price Indexes, since most of the components of these indexes enter indirectly into the price deflator.[14] In addition, the price deflator has quality problems of its own connected with the inclusion of government services. Services of government employees enter into the index at prices which merely reflect the wages and salaries of such employ-

---

[14] To the extent that components of the Consumer and Wholesale Price Indexes having upward biases are used to deflate portions of the GNP, the resulting constant dollar magnitudes are underestimated. When the constant dollar GNP containing such underestimated elements is divided into the current dollar GNP to obtain the implicit price deflator, an upward bias is imparted to the deflator.

ees. Since there seems to be little doubt that increasing mechanization and other factors which have improved efficiency in the private sector of the economy have also had some similar effects in government, the method of handling government activities, by failing to take account of such increases in efficiency as have occurred, has almost certainly imparted a further upward bias to the index.

The magnitude of the upward biases introduced into our major price indexes by inadequate allowances for quality improvement is very difficult to judge. Some writers have suggested that a considerable part of the inflation that we seem to have experienced in the past few years and which has caused so much concern is merely a statistical illusion resulting from upward biases in our price indexes.[15] Others, recognizing that there is quality deterioration in some areas and that the adjustments made by the Bureau of Labor Statistics may, in some instances, overcorrect for quality improvement, conclude that, on balance, quality changes probably cause relatively little distortion in the price indexes.[16] The whole subject is beclouded by a lack of empirical evidence and by semantic and even philosophical disputes about the meaning of the word "quality," which is admittedly difficult to define in an objective manner that does not reflect the value judgments of the particular observer.[17] The present writer is inclined to the belief that there are significant upward biases, particularly in the implicit price deflator and the Consumer Price Index. However, it is impossible to measure these biases or even to prove conclusively that they are present.

---

[15] For an especially strong expression of this view, see Richard Ruggles and Nancy D. Ruggles, "Prices, Costs, Demand, and Output in the United States, 1947–57," in *The Relationship of Prices to Economic Stability and Growth,* Compendium of Papers Submitted by Panelists Appearing before the Joint Economic Committee (Washington: U.S. Government Printing Office, 1958), pp. 297–308. This article dramatized the quality issue by raising a question as to whether an individual given $1,000 to spend and a choice of ordering goods from a 1948 or a 1957 Sears, Roebuck catalog would invariably select the earlier catalog despite the fact that the prices quoted in it were substantially lower. That is, there is a real question whether the prices of 1957 were higher or lower than those of 1948 when appropriate allowance is made for improvements in quality. In 1966, Richard Ruggles again expressed an unqualified view that inadequate adjustment for quality improvements has created an upward bias in the Consumer Price Index. See his testimony in *Government Price Statistics,* Hearings before the Subcommittee on Economic Statistics of the Joint Economic Committee, 89th Congress, 2d Session, May 24, 25, and 26, 1966 (Washington: U.S. Government Printing Office, 1966), pp. 237–66.

[16] This view was well expressed by Jules Backman in his 1966 testimony before the Joint Economic Committee. See *Government Price Statistics, op. cit.,* pp. 25–27. It may be noted that during World War II, a special study concluded that the Consumer Price Index (then called the "Cost of Living Index") understated the increase in prices under the special conditions of wartime—including deterioration of quality, disappearance of cheaper goods, underreporting of prices actually charged, and decrease of special sales. See *The Consumer Price Index: History and Techniques, op. cit.,* pp. 6–7.

[17] The Stigler Report (see footnote 13) expressed the opinion that, in the absence of empirical evidence, it is very difficult to judge the effects of quality changes on the price indexes. However, the report did stress the need for further research and offered some suggestions for improved handling of the quality problem.

OTHER ECONOMIC INDICATORS

In addition to measures of employment, unemployment, and the price level, there are many other statistical series which serve as important "indicators" of business and economic conditions. We shall only refer briefly to a few of the most important of these indicators.

The Board of Governors of the Federal Reserve System prepares and publishes an index of industrial production, which measures changes in the physical volume of output of manufactures and minerals. In addition to the annual index, a monthly index is available which is adjusted for seasonal variation. This index is compiled in a manner analogous to a price index, except that the quantities vary from period to period while unit values (or prices) serve as weights and are held constant. At the present time, the base period of the index is 1957–59 and the weights employed are values of 1957. Subindexes are prepared and published covering a great many individual industries and portions of the manufacturing and mining sectors of the economy.[18] The index measures physical output at all stages in manufacturing and mining industries, including intermediate as well as final products. It does not cover production on farms, in construction industries, public utilities, transportation, or trade and service industries. The manufacturing and mining sectors covered by the index account for roughly one third of national income and are particularly important in the analysis of business fluctuations.

The Department of Commerce compiles extensive information concerning sales and inventories for the manufacturing, wholesale trade, and retail trade sectors of the economy. As we shall see, changes in business inventory stocks have played a key role in business fluctuations, and, accordingly, information concerning inventory and sales movements is exceedingly important in the analysis of business conditions. In addition, information is collected concerning manufacturers' new orders and backlogs of unfilled orders, and this is also of considerable importance.[19]

A great deal of detailed statistical information relating to specific industries and sectors of the economy is collected by both private and governmental sources. These data relate to residential and other construction activity, production and sales of consumer durable goods, rates of producton in specific industries, and so on. In addition, a great many financial statistics relating to the banking system, other financial institutions, and the financing of economic activity, are available.[20]

---

[18] The total index, together with detailed breakdowns, is published each month in the *Federal Reserve Bulletin.*

[19] All of this information appears each month in the *Survey of Current Business.*

[20] The best sources of the statistical materials referred to in this paragraph are the *Survey of Current Business* and the *Federal Reserve Bulletin,* the latter publication being a particularly important source of financial statistics.

PART II

Aggregate Demand and Income Determination

INTRODUCTION

It is useful to analyze the forces causing general movements in national production, employment, and the general price level in terms of relations between demand and supply. In the case of the market for an individual good or service, we expect that if there is excess demand—that is, if the demand exceeds the supply at the prevailing price—output of the product will increase or its price will rise or, more likely, both. On the other hand, if there is excess supply—that is, if supply exceeds demand at the prevailing price—output will be reduced or the price will fall or both. If price and output are such that demand and supply are equated, we say that the market is in equilibrium—meaning that unless something happens to shift either demand or supply, the prevailing output and price will remain unchanged. A similar type or analysis involving *aggregate demand* and *aggregate supply* can be applied to the entire economy, although, as we shall see, it contains some pitfalls which are not present in the case of the market for a single commodity.

As a first approximation, subject to later qualification, we can say that (1) if aggregate demand is greater than aggregate supply at the currently prevailing level of prices, producers will either increase their output of goods and services and their employment of factors of production, or raise their prices, or both; (2) if aggregate demand is less than aggregate supply, producers will either reduce their output and their employment of productive factors, or lower their prices, or both; and (3) if aggregate demand is equal to aggregate supply, output, employment, and prices will remain constant. These possibilities are summarized in the following table:

Condition	*Result*
Aggregate demand > Aggregate supply	Expansion
Aggregate demand = Aggregate supply	Equilibrium
Aggregate demand < Aggregate supply	Contraction

It should be understood that expansion means an increase in GNP in current dollars, which may be attributable to an increase in real output and employment, or in prices, or in both; while contraction means the opposite.

The Components of Aggregate Demand

In an economy without foreign trade, aggregate demand may be defined as the amount of goods and services (final output) demanded by consumers, by business, and by government.[1] That is, the components of aggregate demand are personal consumption expenditures (C), gross private domestic investment (I), and government purchases of goods and services (G). In accordance with the explanation of GNP from the expenditure side in Chapter 2, we have

$$C + I + G = Y$$

where Y is total output (GNP).

As used in this equation, I is *ex post* or *actual* investment, as it turns out to be after the period under consideration (say, a quarter or a year) is completed. For our present purposes, we need to use a different concept of investment, *ex ante* or *planned* investment. This is the amount of investment in plant, equipment, and inventories that businessmen intend to engage in during the period. Planned investment, I_p, may be either smaller or larger than actual investment, I. We may define

$$\text{Aggregate demand} = C + I_p + G$$

Aggregate demand defined in this way may be either greater than, equal to, or less than aggregate supply—that is, total output or GNP.[2] Thus, the conditions of expansion, equilibrium, or contraction may be restated as in the following way:

$$C + I_p + G > Y \qquad \text{Expansion} \qquad (1)$$
$$C + I_p + G = Y \qquad \text{Equilibrium} \qquad (2)$$
$$C + I_p + G < Y \qquad \text{Contraction} \qquad (3)$$

Since Y is the total currently produced output available to satisfy the demands of consumers, business, and government, if aggregate demand

[1] Foreign trade involves some complications which we shall take up in Part V.

[2] The reader will note that we are assuming for simplicity that the spending plans of consumers and government are exactly satisfied. Moreover, as will be seen below, we assume that the government's tax collections turn out as anticipated. Thus, it is only private investment and private saving plans that may not be realized.

exceeds Y, as in condition (1) this demand could only be met by unintended depletion of inventories of existing goods. Under these conditions, if businessmen expect demand to continue at this level, they will speed up their production (or raise their prices) in the next period to bring supply into line with demand (and perhaps to make up for the unexpected and undesired depletion of inventories). Conversely, if aggregate demand is less than Y, as in condition (3), stocks of unsold goods will pile up in the form of unintended inventory investment; and businessmen will reduce their production rates (or lower their prices) in order to bring supply into line with demand (and perhaps to work off the unexpected accumulation of inventories).

Since $C + I + G = Y$ under all conditions, the first condition above, leading to expansion, implies that $I_p > I$, while the third condition, leading to contraction, implies that $I_p < I$. In each of these cases, the amount of unintended or unplanned inventory disinvestment or investment must be just sufficient so that when this is included along with planned investment, as it is in calculating ex post or actual investment (I), the national income accounting identity, $C + I + G = Y$, will always be satisfied.

If condition (2) above is satisfied—that is, if $C + I_p + G = Y$—we may say that the economy is in *production equilibrium;* i.e., that production is adjusted to demand.

Disposition of Income

There is a further complication not taken into account thus far: even if aggregate demand is equal to aggregate supply at a particular time, there may be forces at work that will in due course disrupt this balance. We turn now to a consideration of such forces.

The GNP, designated by Y, can also be viewed as total gross income. All of this income must be used either for consumption, for saving, or to pay taxes. Thus, we have the following identity, derivable from the national income accounts

$$C + S + T = Y$$

Here C is personal consumption expenditures, S is total gross private saving (including personal saving, corporate retained earnings, and capital consumption allowances), and T is all taxes, federal, state, and local.

As in the case of investment, it is useful to distinguish two different concepts of saving. As used in the above equation, S is ex post or actual saving as recorded in the national accounts. We need to use the concept of ex ante or planned saving, S_p, which may be larger or smaller than S.

If income in a period turns out to be larger than was anticipated, actual saving will turn out to be larger than was planned. When this happens, consumption expenditures will increase in the next period as households adjust their saving downward in the direction of the desired level; this will increase aggregate demand, leading to a rise in production and income. Conversely, if income turns out to be smaller than was expected, actual saving will be less than planned saving, and households will reduce consumption later on in an effort to restore their saving to the desired level; this will reduce aggregate demand, leading to a decline in production and income. Thus, we have the following conditions

$C + S_p + T < Y$	Expansion	(4)
$C + S_p + T = Y$	Equilibrium	(5)
$C + S_p + T > Y$	Contraction	(6)

Since $C + S + T = Y$ under all conditions, the first condition above implies $S_p < S$, leading to expansion, while the third condition implies that $S_p > S$, leading to contraction. In each of these cases, there must be just enough unplanned saving or dissaving to make actual saving (S) sufficient to satisfy the national income accounting identity $C + S + T = Y$.

If condition (5) above is satisfied—that is, if $C + S_p + T = Y$—the economy may be said to be in *income equilibrium*. This means that income recipients are satisfied with the disposition of income among consumption, saving, and taxes.

Overall Equilibrium

We can set the left-hand side of (2) equal to the left-hand side of (5), since each of these expressions is equal to Y. This gives us the condition

$$C + I_p + G = C + S_p + T$$

If this condition is satisfied, the economy is in both production equilibrium and income equilibrium. Subtracting $(C + G)$ from both sides of this equation, the condition can be restated as

$$I_p = S_p + (T - G)$$

That is, for the economy to be in overall equilibrium—including both production equilibrium and income equilibrium—it is necessary for ex ante or planned investment to equal ex ante or planned saving plus the surplus in the government budget. In fact, by combining the inequalities (1) and (3) with the inequalities (4) and (6) and subtracting $(C + G)$ from both sides, we can obtain the following set of conditions

Condition	*Result*	
$I_p > S_p + (T - G)$	Expansion	(7)
$I_p = S_p + (T - G)$	Equilibrium	(8)
$I_p < S_p + (T - G)$	Contraction	(9)

If condition (7) holds, aggregate demand must exceed income, or planned saving must be less than actual saving, or both. Any of these circumstances will lead to expansion. Conversely, if condition (9) holds, aggregate demand must be less than income, planned saving must be greater than actual saving, or both. Any of these circumstances will lead to contraction. Thus, the direction of movement of the economy may be said, at least to a first approximation, to depend on the relation between ex ante or planned investment, on the one hand, and ex ante or planned saving plus the government budget surplus, on the other.[3]

Some Complications and Qualifications

There are several points that the reader should especially take note of concerning the above analysis.

1. In the case of a single market, it is customary to assume that the demand is independent of the supply. This assumption is ordinarily approximately true as long as the analysis applies to a single commodity or a relatively small sector of the economy. For the economy as a whole, however, it is distinctly not true, since an increase in output (supply) causes an equal increase in income, thus generating an increase in purchasing power which is likely to lead to a rise in demand. For this reason, the analysis of the economy as a whole is substantially different from the traditional supply and demand analysis of a single market.

2. The condition that if $I_p = S_p + (T - G)$, output, employment, and the price level will be stable is only an approximation. This condition may be satisfied for the economy as a whole under circumstances in which there are excess demands in some markets and excess supplies in other markets as long as the sum of all the excess demands equals the sum of all the excess supplies. Under these conditions, output and employment and/or prices will be rising in the markets having excess demands, and falling in the markets having excess supplies. However, the effects on output, employment, and prices in these different markets may not cancel out; for example, in the markets in which there is excess demand the adjustment may predominantly take the form of increases in prices, while in the markets having excess supply the adjustment may occur mainly through reductions in output and employment with little reduction in prices. Thus, it is possible for output to fall and the average of all prices to rise under circumstances in which $I_p = S_p + (T - G)$. On the other hand, the opposite result—rising output and employment and falling average prices—could also occur, although this result is somewhat less likely. We shall, however, ignore these problems for

[3] For a more extensive discussion of the relation between planned and actual investment and saving, see pp. 147–52 below.

the present and assume that output, employment, and prices will be stable if $I_p = S_p + (T - G)$.[4]

3. There is an asymmetry in the behavior of our economy which may be noted briefly at this point. Prices are generally recognized to be somewhat more flexible in an upward than in a downward direction. Thus, the tendency is for prices to rise when $I_p > S_p + (T - G)$, while output and employment are likely to feel the chief brunt of the adjustment when $I_p < S_p + (T - G)$.

4. The equilibrium which occurs when $I_p = S_p + (T - G)$ may involve considerable unemployment of the factors of production. This is a question to which we shall give considerable attention later on.

Plan of Part II

The purpose of this part of the book is to explain how the conditions of macroeconomic equilibrium, described above, are achieved; to consider how the equilibrium position of the economy is changed as a result of fiscal or monetary policy actions of the government or displacement of certain relationships in the private sector; and to discuss, to some extent, the dynamic process by which the economy moves from one position of static equilibrium to another.

We begin by examining, in Chapter 5, the determinants of personal consumption expenditures as recorded in the national income and product accounts. Emphasis is placed on the response of consumption to changes in disposable personal income, but other factors influencing consumption, including wealth, interest rates, consumer expectations, and so on, are considered.

In Chapter 6 we develop the static multiplier—that is, the change in the equilibrium level of income produced by an autonomous change in expenditures, taking account of the response of consumption induced by the change in disposable personal income. In this chapter, the analysis is extended to take account of changes in tax revenues, transfer payments, corporate profits, and dividends that are induced directly or indirectly by changes in income. Business investment is taken to be exogenous throughout this chapter, the purpose being to analyze the effects of exogenous changes in expenditures on equilibrium GNP after allowing for induced changes in consumer expenditures.

Chapter 7 deals with the time path of movements of income from one equilibrium position to another. In this chapter, as in Chapter 6, business investment is generally taken to be exogenous; however, at the end of the chapter some attention is given to the way in which inventory adjustments induced by changes in sales may affect the time path of income adjustments.

---

[4] This matter will come up again in the discussion of inflation in Chapter 16.

The chapter also contains some further discussion of the relations between ex ante and ex post saving and investment.

Chapters 8, 9, and 10 deal with the determinants of investment, focusing primarily on business expenditures for plant and equipment. In Chapter 8, the marginal efficiency of investment concept is developed. Chapter 9 deals with the dynamics of investment, with emphasis on the acceleration principle, while Chapter 10 discusses the financing of investment in a dynamic setting. In a concluding section of Chapter 10, the analysis reverts to a static setting in order to develop the so-called *IS* curve, showing the various combinations of income and interest rate which will equilibrate the real sector of the economy.

Chapter 11 deals with the demand for money and the determination of the interest rate, while Chapters 12 and 13 draw together much of the analysis developed in Part II. Chapter 14 introduces some further refinements relating primarily to the behavior of the financial sector of the economy.

Chapter 5 THE DETERMINANTS OF CONSUMER EXPENDITURES

Personal consumption expenditures constitute by far the largest of the major components of aggregate demand for goods and services, amounting consistently to over 60 percent of GNP in recent years. The sheer magnitude of consumption spending indicates that changes in such spending may have profound effects on economic conditions. Accordingly, in any analysis of the determination of national income and employment, careful attention must be given to the factors influencing consumer spending.

Consumption as a Function of Income

The proposition that consumption expenditures will ordinarily depend chiefly on the level of income—rising as incomes rise but by less than the rise in income—was first advanced over 30 years ago by John Maynard Keynes.[1] Since that time, a great deal of research has been done on the subject of individual and aggregate consumption. As a result of this research, Keynes's insight has been refined and modified in substantial fashion; nevertheless the notion that income is, under most circumstances, the *main* systematic determinant of consumption is still generally accepted.

During cyclical fluctuations in economic activity, the relation between consumption and income is reasonably well depicted by a function of the type shown in Figure 5–1. This function is of the *linear* form (i.e., when plotted on graph paper it becomes a straight line), which can be expressed by an equation of the type

$$C = cY + C_0$$

In this equation, C is aggregate consumption expenditure and Y is aggregate income. The equation says that if income were zero, consumption would be C_0, and that as income changes, consumption changes by $100c$ percent of the change in income. The specific consumption function which is shown in Figure 5–1 is

[1] J. M. Keynes, *The General Theory of Employment, Interest, and Money* (New York: Harcourt, Brace & Co., 1936), chap. 8.

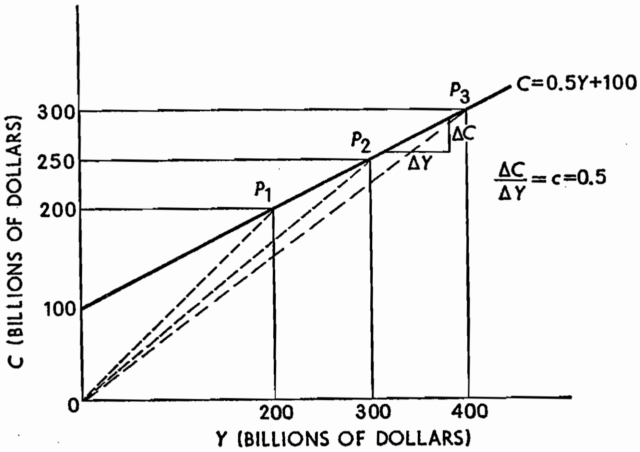
$$C = 0.5Y + 100$$

Here C and Y are expressed in billions of dollars, and the equation states that if income were zero, consumption would be $100 billion, and that when income changes, consumption will change in the same direction by 0.5 (or 50 percent) of the change in income.[2]

The slope of the consumption function—c in the general case and 0.5 (or 50 percent) in the specific example shown in Figure 5–1—is called the

FIGURE 5–1

Short-Run Consumption Function



marginal propensity to consume. Thus the marginal propensity to consume may be defined as the *change in consumption expenditures divided by the change in income that causes it.* For a linear consumption function, the marginal propensity to consume is constant—that is, it does not change with income. However, the *average propensity to consume,* which is simply *the proportion of income spent on consumption* declines as income rises. In the example we are using (see Figure 5–1), if income is $200 billion, consumption will be $200 billion, or the average propensity to consume will be 100 percent; if income is $300 billion, consumption will be $250 billion, or the

[2] The meaning of the constant C_0 ($100 billion in this example) is often puzzling to students. In one sense it is the amount of consumption spending that would occur if income were zero. However, income could scarcely be zero except under conditions of utter chaos, in which case the normal consumption function would hardly apply. Actually, therefore, C_0 has no *economic* meaning. It serves the purely *algebraic* function of helping to designate a particular straight line. One way to specify a straight line is to give the value at which the line intersects the vertical axis, together with the slope of the line. In this case the line intersects the vertical (consumption) axis at a value of C_0 and has a slope of c.

average propensity to consume will be 83 percent; and if income is $400 billion, consumption will be $300 billion, or the average propensity to consume will be 75 percent. Thus, even though the marginal propensity to consume is constant, the proportion of income consumed falls as income rises.

What Income Concept? From the discussion in Chapter 2, it seems apparent that consumption expenditures should be more closely related to disposable personal income, which is the amount of income left to consumers after paying their taxes, than to national income, gross national product, or personal income. This is true, and we shall accordingly use disposable personal income in our analysis.

Real or Money Values? Another question that arises is whether the consumption and income data should be deflated and expressed in constant dollars or whether they should be valued at current prices. Since it is plausible to assume the absence of "money illusion"—i.e., that if everyone's income were, for example, doubled and at the same time all prices were doubled, the volume of consumer goods purchased would be unchanged—it seems preferable to deflate both income and consumption expenditures by use of an index of consumer prices.[3] It usually turns out that the relationship between income and consumption is improved when this is done

Aggregate or per Capita Values? Finally the question arises as to whether the relationship is between aggregate income and aggregate consumption or whether it is between per capita income (i.e., aggregate income divided by the population) and per capita consumption. Since if income remains constant while population, for example, doubles, the average person will be only half as well off as before and will presumably adjust his consumption accordingly, it is better to relate per capita income and per capita consumption.

---

[3] Since both consumption expenditures and income are ordinarily deflated with the same index number (of consumer prices), students are often puzzled as to why deflation makes any difference. Suppose income and consumption values at current prices are $108 billion and $86.4 billion, respectively, in one year, and $144 billion and $108 billion, respectively, in the next year. The marginal propensity to consume computed from these data would be

$$\frac{\Delta C}{\Delta Y} = \frac{108.0 - 86.4}{144.0 - 108.0} = \frac{21.6}{36.0} = 0.6$$

Now suppose the price index (on some third year as a base, of course) was 108 in the first year and 120 in second year. The marginal propensity to consume computed from the deflated data would be

$$\frac{\Delta C}{\Delta Y} = \frac{90 - 80}{120 - 100} = \frac{10}{20} = 0.5$$

As a matter of fact, deflation will always change the consumption function except when consumption is proportional to income (i.e., the marginal propensity to consume is equal to the average so that the function is of the type shown in Figure 5–3).

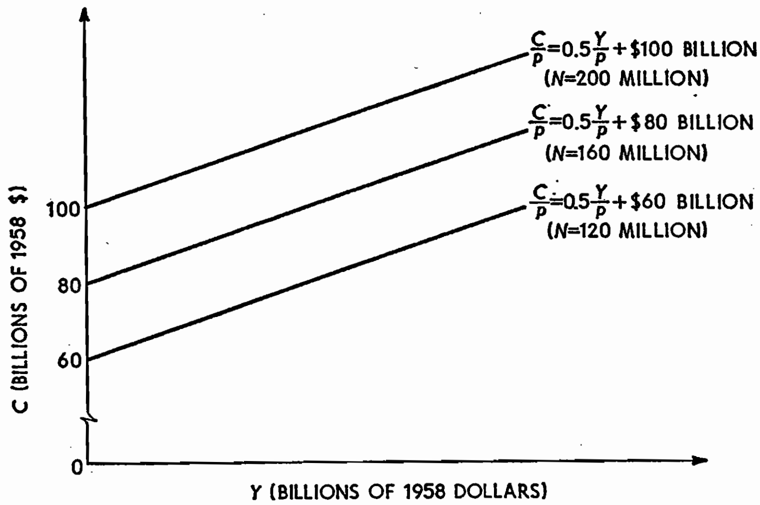
Summary. Thus, the relationship we are concerned with is best formulated as follows:

$$\frac{C}{Np} = c\frac{Y}{Np} + c_0$$

where N is population, p is the consumer price index, Y is aggregate disposable personal income valued at current prices, and c_0 is per capita real

FIGURE 5-2

Shifts in Aggregate Consumption Caused by Changes in Population



Per capita consumption function (in 1958 dollars):

$$\frac{C}{Np} = 0.5\frac{Y}{Np} + 500$$

consumption when per capita real disposable personal income is zero. This can easily be converted into an aggregate consumption function by multiplying through by population (N); thus

$$\frac{C}{p} = c\frac{Y}{p} + c_0N$$

Thus, C_0 (which was equal to \$100 billion in the example used above) becomes c_0N, and, other things remaining constant, an increase in the population causes the consumption function to rise. This is a matter of some importance, and it is illustrated in Figure 5–2.

High Employment versus Cyclical Consumption Functions

During cyclical fluctuations in which income first falls below the full employment level and then rises again toward such a level, a consumption function of the type shown in Figures 5–1 and 5–2—i.e., one which is characterized by a constant marginal propensity to consume and an average propensity to consume which declines as income rises—seems to provide a reasonably satisfactory explanation of the behavior of consumption expenditures. In the course of the prolonged decline in economic activity from 1929 to 1933, the drop in real per capita personal consumption expenditures was 72.3 percent of the decline in real per capita disposable personal income, and during the ensuing recovery period from 1933 to 1937 the corresponding percentage was 72.4 percent. During the postwar period, the American economy has passed through four relatively mild recessions—those of 1949, 1953–54, 1957–58, and 1960–61. While there has been considerable variation in the behavior of consumption expenditures from one recession to another, due presumably to the operation of factors other than income affecting consumption, as discussed below, the marginal propensity to consume during the expansion and contraction phases of the four postwar recessions has averaged approximately 70 percent. That is, the decline in real per capita personal consumption expenditures has averaged about 70 percent of the decline in real per capita disposable personal income during the contraction phases of the cycles, while, on the same basis, the rise in consumption has averaged about 70 percent of the rise in income during recovery phases prior to the time reasonably full employment has been restored.

On the other hand, when periods of high-level employment and prosperity are compared, the *average* propensity to consume is remarkably constant in the neighborhood of 91 to 93 percent. For example, in the years 1952, 1953, 1956, 1965, and 1968—all years of high employment and considerable prosperity—the ratio of consumption to income ranged only from 90.9 to 91.7 percent. This was true in spite of the fact that disposable personal income in 1958 dollars rose substantially during this period—from $1,678 per capita in 1952 to $2,474 per capita in 1968, an increase of 47.4 percent. In fact, even going back some 30 years to the prosperity year 1929, we find that consumption was 92.7 percent of disposable personal income.[4] This suggests that during periods of approximate full employment, the consumption function is of the type

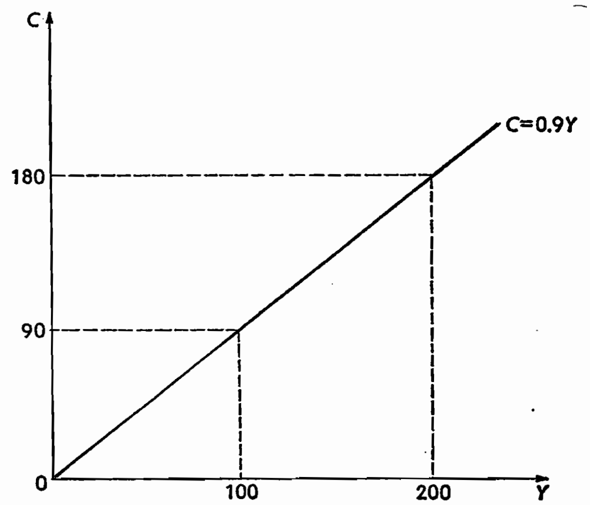
$$C = 0.9Y$$

[4] The constancy of the average propensity to consume (or save) over long periods of time was first called to the attention of economists by Simon Kuznets. See his study, *Uses of National Income in Peace and War,* Occasional Paper No. 6, National Bureau of Economic Research, 1942.

i.e., that both the average and marginal propensities to consume are equal to approximately 90 percent. That is, the consumption function is well depicted by a straight line passing through the origin of the consumption and income axes and having a slope of 0.9, as shown in Figure 5–3.

Thus, past behavior of the economy indicates that the marginal propensity to consume out of disposable personal income is about 70 percent during periods of cyclical fluctuations and about 90 percent during periods of steady growth at full employment. At first glance, this seems an odd and rather

FIGURE 5–3
Consumption Function with Growing Income under Conditions of Full Employment



puzzling relationship. However, several possible explanations have been suggested by students of consumption.

One explanation rests on the so-called *relative income hypothesis.* According to this hypothesis, due to the psychology of "keeping up with the Jones," the fraction of income consumed by an American family commonly depends not so much on the absolute level of the family's income as upon the relative position that it occupies in the income distribution. This being the case, it is possible that during periods of full employment and generally rising living standards, when nearly everyone's income is rising and there are few changes in the relative positions of families in the income distribution, consumption spending is "ratcheted" upward in such a way that the ratio of consumption to income (the average propensity to consume) remains constant.[5] On the other hand, when a recession occurs and incomes fall, the higher incomes of the past will have set living standards which consumers desire to maintain (and in any case find difficult to give up immediately due to contractual commitments). Accordingly, as incomes fall, consuming units resist reductions in their consumption standards and tend to reduce saving instead, so that consumption falls less than in proportion to income. During recovery—prior to the time the prerecession level of income is restored—consumption moves back along the path followed during the decline in income, as consuming units seek to restore their saving to its normal level.[6]

According to this view, consumption spending depends partly on current income and partly on living standards which are based in turn on the highest level of income reached in the past. Suppose, for example, that the relationship is[7]

$$c = 0.7y + 0.2y_p \qquad (1)$$

where c is per capita real personal consumption expenditures, y is the current year's per capita real disposable personal income, and y_p (for "peak income") is the highest level of disposable personal income ever reached in any past year. Thus, consumption is equal to 70 percent of the current year's income plus 20 percent of past peak income.

In order to show how such a relationship could explain the simultaneous existence of a short-run marginal propensity to consume of 70 percent and a long-run marginal (and average) propensity to consume of about 90 percent, let us first consider a situation in which the economy is growing steadily at a rate of $100g$ percent per year. In such a situation, the past peak income for any year is obviously the income of the year before; in fact we have

$$y = (1 + g)y_p$$

---

[5] The point can be illustrated by means of a simple example. To begin with, suppose there are three families with incomes of \$2,000, \$4,000, and \$8,000, consuming \$2,000, \$3,600, and \$5,600, or 100 percent, 90 percent, and 70 percent of income, respectively. Thus, total income is \$14,000, and total consumption is \$11,200 or 80 percent of income. Now, according to the *relative income hypothesis,* if all three families' incomes double (rising to \$4,000, \$8,000, and \$16,000, respectively), their average propensities to consume, which depend on each family's income *relative* to the others, will not change (there having been no changes in relative position), total consumption will rise to \$22,400 out of total income of \$28,000, and the average propensity to consume of the society as a whole will remain at 80 percent.

[6] For a detailed exposition of the argument summarized in this paragraph, see J. S. Duesenberry, *Income, Saving, and the Theory of Consumer Behavior* (Cambridge: Harvard University Press, 1952).

[7] This is basically the same type of equation as that used in Franco Modigliani, "Fluctuations in the Savings-Income Ratio: A Problem in Economic Forecasting," in *Studies in Income and Wealth,* Vol. 11, National Bureau of Economic Research, 1949. A similar equation, although differing in detail, is developed in Duesenberry, *op. cit.*

or

$$y_p = \frac{y}{1+g}$$

Substituting this for y_p in Equation 1, we have

$$c = 0.7y + \frac{0.2}{1+g}y$$

or

$$c = \left(0.7 + \frac{0.2}{1+g}\right)y$$

Over the past 20 years or so the average growth of real per capita disposable personal income has been about 2.25 percent per year (i.e., $g = 0.0225$). In this case, $0.2/(1 + g)$ is equal to $0.2/1.0225$, or 0.196, and we have

$$c = 0.896y$$

Thus, the marginal (and average) propensity to consume under such conditions of steady growth is approximately 90 percent.[8]

Suppose now that a recession occurs and income falls below the full employment level. For purposes of illustration, suppose that the highest real per capita disposable personal income attained before the recession was \$1,900. Thus, during the decline and recovery y_p will remain constant at \$1,900, and according to Equation 1, the consumption function will be

$$c = 0.7y + 0.2(1{,}900)$$

or

$$c = 0.7y + 380$$

The results of our analysis are summarized graphically in Figure 5–4. The line *OD* is the relation $c = 0.9y$ and expresses the path along which consumption moves in relation to income when income is growing at a rate of around 2.25 percent. The lines *AA′*, *BB′*, and *CC′* are successive cyclical consumption functions for three business cycles for which peak income is assumed to be \$1,900 per capita, \$2,200 per capita, and \$2,500 per capita, respectively. Suppose we start at point *P* on line *OD* and income grows steadily under conditions of full employment until we come to *A*, at which point for some reason a decline in business activity and income sets in. Instead of moving back over the path *OD* along which it has just come

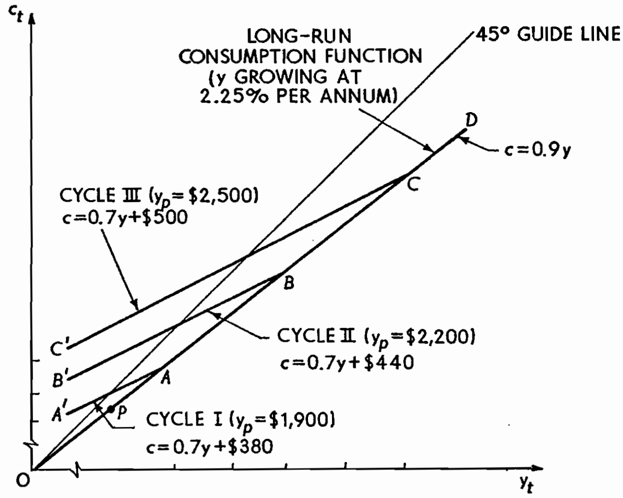
[8] It may be noted that the average propensity to consume under conditions of full employment depends upon the rate of growth of income. However, for plausible rates of growth of from, say, 1 to 5 percent ($g = 0.01$ to 0.05), the average propensity to consume will not depart very far from 90 percent. We are not, of course, trying to make precise calculations, but merely to establish approximate orders of magnitude.

consumption will fall back along the path AA' as income falls and then move back to point A (again following AA') as recovery develops.[9] When income gets back to the full employment level at A, it once again takes the full employment path, with consumption and income moving along OD until point B is reached and another cyclical movement occurs. And so on.

Starting from the consumption function formulated in per capita terms (Equation 1 above), it is possible to translate into aggregative terms by multiplying through by the population (designated by N). If C and Y are aggregate consumption and aggregate income, respectively, while c and y

FIGURE 5–4

Hypothetical Cyclical and Long-Run Consumption Functions



are the corresponding per capita magnitudes, we have $C = Nc$ and $Y = Ny$. Accordingly, when we multiply both sides of Equation 1 by N, we obtain

$$C = 0.7Y + 0.2y_pN$$

Since the population in 1968 was approximately 201.2 million and per capita disposable personal income in 1968 (the highest level reached at that time) was $2,474 (valued at 1958 prices), the short-run consumption function in aggregative terms for 1968 would be approximately

$$C = 0.7Y + (0.2)(\$2{,}474)(201.2 \text{ million})$$

[9] The results obtained by Duesenberry (*op. cit.*) are quite similar to those given here (which follow Modigliani), except that Duesenberry's short-run consumption functions are curved lines (nonlinear) rather than straight lines such as those shown in Figure 5–4.

or

$$C = 0.7Y + 100 \text{ (in billions of 1958 dollars)}$$

This may be taken as a rough approximation of the path that might have been followed by consumption if a decline in economic activity were to have occurred in 1968 with population remaining constant.

In addition to the analysis based on the relative income hypothesis, there are two other ways of explaining and reconciling the difference between the long-run and cyclical behavior of consumption. One is the so-called *life cycle hypothesis*,[10] which begins from the utility functions of individuals (or households) and, by making certain assumptions about these utility functions, derives the proposition that

$$c_t = kv_t$$

where c_t is the consumption of an individual, v_t is the present value of the total resources accruing to the individual over the rest of his life, and k is a factor of proportionality. The total resources becoming available over the individual's life can be taken as the sum of the individual's net worth at the end of the previous period plus his nonproperty income during the current period plus the sum of the discounted values of the amounts of nonproperty income he expects to receive in future periods.

Assuming that the discounted sum of expected future nonproperty income is proportional to current nonproperty income and aggregating over individuals, a total consumption function can be derived in the following form

$$C_t = aY_t^n + bA_{t-1} \tag{2}$$

where Y_t^n is aggregate nonproperty income in period t and A_{t-1} is aggregate net worth at the end of period $t - 1$. Property income in period t is rA_{t-1} where r is the rate of return on net worth. Total income (Y) is equal to nonproperty income plus property income; that is,

$$Y_t = Y_t^n + rA_{t-1}$$

or

$$Y_t^n = Y_t - rA_{t-1} \tag{3}$$

Substituting Equation 3 into Equation 2 and simplifying, we obtain

$$C_t = aY_t + (b - ar)A_{t-1}$$

Statistical estimates of Equation 2 suggest that the values of a and b are about 0.7 and 0.06, respectively. If we take r, the rate of return on net worth to be constant at about 0.04, this gives us approximately

[10] For a detailed discussion, see Albert Ando and Franco Modigliani, "The 'Life Cycle' Hypothesis of Saving: Aggregate Implications and Tests," *American Economic Review*, Vol. LIII, March 1963, pp. 55–84.

$$C_t = 0.7Y_t + 0.03A_{t-1} \quad (4)$$

To see the long-run implications of this consumption function, divide both sides of Equation 4 by Y_t.[11] This yields

$$\frac{C_t}{Y_t} = 0.7 + 0.03\frac{A_{t-1}}{Y_t}$$

In the long run, income and net worth can be expected to grow at the same rate, and empirical evidence indicates that the ratio of net worth to income has been in the neighborhood of 5. Substituting a value of 5 for A_{t-1}/Y_t, we obtain

$$\frac{C_t}{Y_t} = 0.85$$

That is, on these assumptions, in the long run consumption will be a constant proportion of income, the proportion being 85 percent. Thus, the long-run consumption function will be a straight line like *OC* in Figure 5–3.

In the course of a business cycle, investment will decline and the rate of growth of net worth will be reduced. If, as an approximation, we assume that it stops growing entirely and is constant at a level of $3,000 billion, the short-run consumption function for that cycle would become (from Equation 4),

$$C_t = 0.7Y_t + 90$$

This consumption function would be a straight line with an intercept of $90 billion on the vertical axis, similar to the function depicted in Figure 5–1. Combining short-run and long-run behavior, we would get a picture similar to that depicted in Figure 5–4.

The third approach to a reconciliation of the secular and cyclical behavior patterns of consumption is the *permanent income hypothesis* developed by Friedman.[12] Both income and consumption are divided into "permanent" and "transitory" components. Permanent consumption (C_p) is taken to be proportional to permanent income (Y_p); that is,

$$C_p = kY_p$$

The ratio, *k*, is not, in principle, a constant but rather would vary with the interest rate at which consumers could borrow or lend, the age structure of the population, and so on, although Friedman concludes that it has, in fact, been roughly constant over a long period of time. In addition, we have

$$Y = Y_p + Y_t$$
$$C = C_p + C_t$$

[11] For a more detailed derivation of the estimates given here, see *ibid.*, pp. 76–80.

[12] See Milton Friedman, *A Theory of the Consumption Function* (Princeton, N.J.: Princeton University Press, 1957).

where Y and C are measured income and consumption (as recorded in the national income accounts, for example), Y_p and C_p are permanent income and consumption, and Y_t and C_t are transitory income and consumption. The key assumption is that there is no systematic relation between transitory consumption and transitory income—that is, in effect, the marginal propensity to consume out of transitory income is zero. Transitory income would include "windfalls," such as lottery winnings, temporary variations in income due to seasonal factors, cyclical fluctuations, etc., while permanent income would be the income judged by the individual to be "normal."

In essence, permanent income is a concept that is oriented toward the future. However, there is no practical way of measuring or quantifying income in this sense. Consequently, in his work with time series Friedman takes permanent income to be a moving average of the incomes of past years. The moving average covers a period of 17 years, with the current year's income receiving a weight of 33 percent, the previous year's income a weight of 22 percent, and the preceding 15 years' incomes a series of rapidly declining weights. For the nonwar years from 1905 through 1951, Friedman obtains the consumption function

$$C = 0.88Y_p$$

That is, the marginal (and average) propensity to consume out of permanent income is 88 percent. This can be viewed as the long-run consumption function corresponding to line OC in Figure 5–3. The short-run marginal propensity to consume, on the other hand, is very low. If the current year's income should fall by \$100, this would reduce permanent income by only about \$33, since, as indicated above, the current year's income has a weight of 33 percent in the computation of permanent income. With a marginal propensity to consume out of permanent income of 88 percent, this would cause consumption to decline by about \$29 ($33 \times 0.88 = 29.04$). That is, the one-year marginal propensity to consume is about 30 percent. This is somewhat misleading, however, since under Friedman's definitions, consumption does not include purchases of consumer durable goods, such as automobiles, furniture, and home appliances. Instead these are treated as investment and therefore constitute one of the forms that consumer saving may take. Thus, the use of a windfall of income to buy durable goods would not be inconsistent with Friedman's assumption that the marginal propensity to consume out of transitory income is zero. Similarly, in the case of the aggregate consumption function referred to above, the 30 percent one-year marginal propensity to consume refers only to expenditures for nondurable goods and services. In addition, there might be a substantial adjustment in purchases of durable goods, making the short-run response of consumption to income comparable to that uncovered by other studies if the same definitions were employed.

The consumer unit's wealth may be viewed as the present value of its expected future income receipts discounted to the present. Or, to turn the relation around, its permanent income may be viewed as the current year's return on its wealth in this sense. That is,

$$Y_p = iW$$

where Y_p is permanent income, W is wealth (including "human" wealth, that is, the present value of the expected future income from work effort), and i is the appropriate rate of return on wealth. When the permanent income hypothesis is viewed in this way, it clearly bears a close relation to the life cycle hypothesis discussed above.

Factors Other than Income Affecting Consumption[13]

Although income and perhaps wealth are the most important variables affecting consumption, there are changes in consumer expenditures which can hardly be accounted for by changes in income and wealth alone. We shall consider briefly several other factors which appear to affect consumption.

1. Consumer Attitudes and Expectations.[14] Consumption expenditures are determined not only by *ability* to buy, which depends to a considerable extent upon income, but also by *willingness* to buy. The latter undergoes variations which are related to changes in consumer attitudes and expectations. As far as *income expectations* are concerned, there is evidence that, other things being equal, those who expect their incomes to increase have a tendency to spend more and save less than do those who expect their incomes to decline. Thus, we may expect that when the general business outlook is favorable, the consumption function may be higher than when it is unfavorable. In particular, expectations of higher incomes seem to stimulate more frequent and larger purchases of "big ticket" durable goods, such as automobiles, electric appliances, and furniture.

Studies of consumer behavior indicate that the effects of *price expectations* are not always what one might expect. For example, one might suppose that expectations of falling prices would cause consumers to postpone spending and wait for the expected declines to materialize; however, there are indications that people may buy more when they expect prices to fall because they associate falling prices with healthy economic conditions.

---

[13] For an excellent discussion of many of the points taken up in this section, as well as further references to the relevant literature, the student is referred to D. B. Suits, "The Determinants of Consumer Expenditure: A Review of Present Knowledge," in *Impacts of Monetary Policy,* Research Studies Prepared for the Commission on Money and Credit (Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1963), pp. 1–57.

[14] The Survey Research Center at the University of Michigan has pioneered in the study of consumer attitudes and expectations. The results are summarized in two books by George Katona entitled *Psychological Analysis of Economic Behavior* (New York: McGraw-Hill Book Co., 1951) and *The Powerful Consumer* (New York: McGraw-Hill Book Co., 1960).

Studies also indicate that price expectations are quite unstable and change frequently and that the reasons for holding price expectations are often quite superficial. Nevertheless, it seems clear that there are times when strong expectations of rising prices may induce a sharp increase in consumer spending. In fact, this seems to have happened at the time of the outbreak of the Korean War in mid-1950.

*2. **Interest Rates and Credit Conditions.*** It used to be felt that saving was markedly influenced by interest rates, although there has always been some uncertainty as to the nature of the effect. Higher interest rates may cause some people to save more, while for others—for example, those who are saving to accumulate a given retirement income—higher interest rates may mean they can achieve their objectives with less saving. Probably the consensus today among economists is that interest rates have no more than a very minor effect on the level of saving. Much saving is apparently motivated by a desire to accumulate funds for some specific future use, such as making a down payment on a house or putting children through college, rather than for income purposes.[15] Saving of this kind is not likely to be much influenced by interest rates. However, *relative* interest rates on different types of claims—such as savings deposits, shares in savings and loan associations, and United States savings bonds—may affect the *form* in which savings are held.

There is one way in which saving and consumption, as measured in the national income accounts, may be influenced by interest rates and credit conditions. This is through the use of *consumer installment credit.* Higher interest rates and tighter credit conditions may restrict consumer expenditures on durable goods financed by installment credit. Thus, credit conditions, while not especially affecting positive saving, may encourage or discourage dissaving through this particular channel.[16]

*3. **Liquid Assets.*** As indicated above, the life cycle hypothesis and, in a sense, the permanent income hypothesis introduce wealth or consumer net worth directly into the aggregate consumption function. However, some studies have placed special emphasis on the accumulated stock of consumer *liquid assets*—demand deposits, currency, time deposits, United States savings bonds, etc.—as a determinant of consumption.[17] The evidence concerning the effects of liquid assets is somewhat mixed. Under normal conditions, the changes that occur in the stock of liquid assets during a year's time are probably not large enough to exert a major influence. Moreover,

---

[15] On this, see Katona, *The Powerful Consumer, op. cit.,* chap. 7, and pp. 222–24.

[16] A recent study has uncovered evidence that consumer demand for durable goods is influenced by interest rates. See M. J. Hamburger, "Interest Rates and the Demand for Consumer Durable Goods," *American Economic Review,* Vol. LVII, December 1967, pp. 1131–53.

[17] See, for example, Arnold Zellner, "The Short Run Consumption Function," *Econometrica,* Vol. XXV, October 1957, pp. 552–67.

changes in household holdings of liquid assets can, in general, scarcely be regarded as exogenous to the system since, like consumption and saving, they are largely subject to the discretion of households. Consequently, their introduction would require the development of a systematic analysis of their underlying determinants. Under some special conditions, however, changes in consumer liquid assets may be especially important and may also be regarded, with reasonable propriety, as being essentially exogenous. For example, immediately after World War II, there was a large upsurge in consumer spending, which was in all probability partly caused by the huge increase in consumer liquid assets that resulted from the methods used to finance the war.[18]

*4. **Distribution of Income.*** Two different aspects of income distribution are often considered by economists: the *personal distribution of income* among individuals or families in different income brackets, and the *functional distribution of income* among the categories of wages, rent, interest, and profits. One might expect that consumption out of a given aggregate income would be smaller the more unequal was the personal distribution of income, because persons with high incomes would save a larger proportion of their incomes than would persons with low incomes. However, the statistical evidence in support of this expectation is not very impressive. Similarly, one might expect that a change in the functional distribution of income away from wages and toward profits would reduce consumption because the saving tendencies of profit recipients would be greater than those of wage earners. There are two reasons why this might be so: first, profit recipients are likely to be in higher income brackets than wage earners and, second, profit recipients have a special motivation to save in order to be able to finance further investment in their businesses. There is indeed some statistical evidence that the propensity to save of profit recipients is greater than that of wage earners. It should be noted, however, that the functional distribution of income bears a reasonably systematic relation to the level of income, so that the effects on consumption of "normal" changes in the functional distribution of income are presumably captured to a considerable extent by aggregate consumption functions of the type discussed above.

*5. **Other Factors.*** Shifts in the consumption function may occur due to changes in consumer tastes, changes in the attractiveness of products, advertising outlays, etc. For example, the heavy demand for automobiles in 1955 and the reduced demand in 1956, which were large enough to affect the overall level of consumption significantly, probably can be attributed largely to factors of this kind. Moreover, the demand for consumer durable goods

[18] At a more abstract level, it is sometimes postulated that consumption spending is influenced by the accumulation of "net claims against the government." The aggregate change in such net claims during any time period is determined by the federal government budget deficit or surplus, and it may therefore properly be regarded as exogenous. For a discussion of the theoretical implications of this idea, see Chapter 15.

has special characteristics, in addition to the use of installment credit referred to above, which cause it to behave differently from the demand for nondurable goods and services.[19] Capital gains or losses resulting from changes in the value of existing wealth are not included in income as measured in the national accounts. At times such gains or losses may exert an appreciable effect on consumption expenditures. Finally, it may be noted that over very short periods of time, such as a quarter of a year, there is a good deal of random variation in consumption, so that the relation between consumption and income for such periods is less reliable than for longer periods.

Method of Handling Determinants Other than Income. Most of the factors discussed above are rather difficult to integrate directly into a model of the economy, because they are frequently caused by changes in consumer psychology or in public policy. Accordingly, we shall view consumption as dependent chiefly on income, in accordance with our earlier analysis. When we wish to take account of the effects of changes in consumer expectations, growth of consumer liquid assets, or other factors, we shall handle such changes by introducing appropriate shifts in the consumption-income relationship.

[19] See Chapter 9, p. 190.

Chapter 6
INCOME DETERMINATION WITH AUTONOMOUS INVESTMENT

We are now ready to consider the determination of equilibrium income in the simple case in which the only types of expenditures are consumption expenditures and investment expenditures. We shall assume that consumption depends upon income in accordance with a short-run consumption function of the type discussed above, and, in order to simplify matters, we shall also assume, for the time being, that all incomes earned are paid out to the factors of production so that disposable personal income and GNP are identical. For the present, we shall not concern ourselves with the determinants of investment but shall take the rate of investment expenditures as given.

As pointed out earlier in this chapter, increases in aggregate demand may raise the level of output and employment, they may merely push up prices, or they may produce a mixture of output-employment increases and price increases; similarly a fall in aggregate demand may reduce output and employment or prices of both. This is a matter to which we shall give considerable attention later on,[1] but for the present we shall avoid dealing with it by assuming that there is ample excess labor and plant capacity so that output and employment can increase without driving up the price level; that is, we shall assume that prices and wages remain constant as output increases, so that money income and real income move precisely parallel to one another.

Let us begin with a simple numerical example. Suppose the consumption function is

$$C = 0.5Y + 120$$

where C and Y are the annual rates of consumption expenditures and income (GNP), respectively, expressed in billions of dollars (since we are assuming prices remain constant, it doesn't matter whether we calculate values at current prices or constant prices). Suppose further that private

[1] See Chapters 15 and 16.

investment expenditures (I) are running at an annual rate of $100 billion per year; that is

$$I = 100$$

Since for the time being we are assuming that there is no governmental activity, aggregate demand is simply equal to consumption expenditures plus investment expenditures. On the other hand, aggregate supply is equal to the amount of output being produced, or alternatively—since the production of a dollar's worth of gross output necessarily generates a dollar of gross income—total gross income, Y. Equilibrium requires that aggregate demand be equal to aggregate supply; thus, the condition for equilibrium is

$$Y = C + I$$

It is a simple matter to find equilibrium income algebraically. We simply substitute $0.5Y + 120$ for C and 100 for I, thus obtaining

$$Y = 0.5Y + 120 + 100$$

and solve for Y, which gives us

$$Y = 440$$

If income is $440 billion, we have:

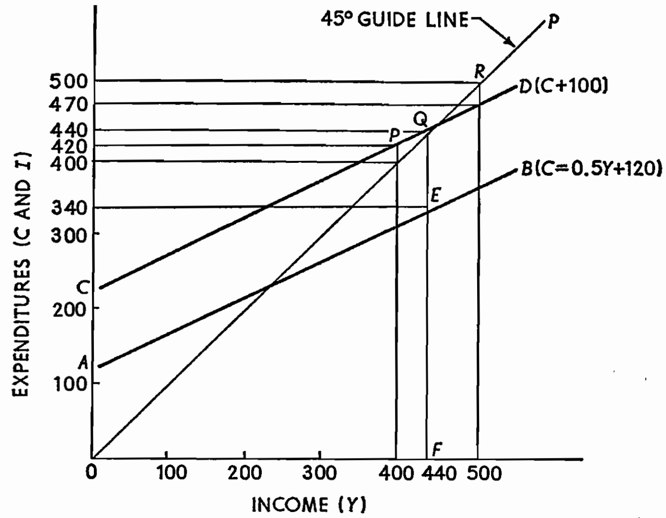
Y (= aggregate supply).............	$440 billion
C (= $0.5Y + 120$)..................	340 billion
I..................................	100 billion
S (= $Y - C$)......................	100 billion
Aggregate demand (= $C + I$)........	440 billion

It is apparent that aggregate demand equals aggregate supply, or, alternately, that planned saving (S) equals planned investment. These two conditions were shown earlier to be alternative ways of specifying aggregate economic equilibrium, or conditions under which income and employment would remain constant.

The equilibrium level of income is depicted graphically in Figure 6–1. In this diagram, expenditures (i.e., consumption and investment) are measured on the vertical axis, while income is measured on the horizontal axis. The line OP, drawn at an angle of 45° between the two axes, has the characteristic that at all points on it expenditures are equal to income. The line AB is simply the graph of the consumption function $C = 0.5Y + 120$ and shows how consumption behaves as income changes. The line CD, which is vertically above AB by $100 billion, shows how total expenditures (consumption and investment together) behave as income changes. The equilibrium position is the place where the total expenditure line, CD, intersects the 45° guideline, since at this point (point Q in Figure 6–1), expenditures

FIGURE 6-1

Determination of Equilibrium Income



are equal to income, or alternatively saving equals investment—investment being the distance (QE) between line AB and the 45° guideline.

It may be noted that in this case the equilibrium is *stable;* that is, if some force should push the economy away from equilibrium it would tend to bounce back toward the equilibrium position. This is illustrated by the following table (billions of dollars):

	Case of Excess Demand	*Case of Excess Supply*
Y (= aggregate supply)	400	500
C (= $0.5Y + 120$)	320	370
I	100	100
S (= $Y - C$)	80	130
Excess demand (= $C + I - Y$)	20	−30
$I - S$	20	−30

In the case illustrated in the first column, income is at the $400 billion level, but there is excess demand of $20 billion—or, alternatively, planned investment exceeds planned saving by $20 billion. Producers will adjust production rates upward to satisfy the excess demand, and income will rise. On the other hand, in the case shown in the second column, income is $500 billion, but there is excess supply (negative excess demand) of $30 billion; or, planned saving exceeds planned investment by $30 billion. Producers, finding that they cannot sell as much as they are producing, will reduce their production, and income will fall. Whenever income is below equilibrium ($440 billion), it will rise, and whenever it is above equilibrium, it will fall: that is, it always tends to move toward equilibrium. (The two cases in the table are shown at points P and R in Figure 6–1.)

The general algebraic presentation of this model, together with its solution, is as follows:

$$C = cY + C_0$$
$$Y = C + I$$

The equilibrium solution is simply

$$Y = \frac{C_0 + I}{1 - c} \tag{1}$$

so that for each value of investment (I), there is an equilibrium level of income (Y). As will be explained in Chapter 7, the model will be stable if the marginal propensity to consume (c) is smaller than unity.

THE MULTIPLIER

Suppose now that for some reason, as yet unexplained, the rate of expenditure on investment increases. What effect will this have on the level of income? Suppose we designate the initial income as Y_1, corresponding to investment of I_1; what we want to know is the level, Y_2, to which income will rise if investment rises to a higher level, I_2. By Equation 1 above, we have

$$Y_1 = \frac{C_0 + I_1}{1 - c} \tag{2}$$

and

$$Y_2 = \frac{C_0 + I_2}{1 - c} \tag{3}$$

Subtracting Equation 2 from 3, we have

$$Y_2 - Y_1 = \frac{C_0 + I_2}{1 - c} - \frac{C_0 + I_1}{1 - c}$$

or, since $\Delta Y = Y_2 - Y_1$ and $\Delta I = I_2 - I_1$, we have

$$\Delta Y = \frac{1}{1 - c}\Delta I$$

The multiplier may be defined as "the change in income divided by the change in investment that caused it"—that is, the multiplier is $\Delta Y/\Delta I$. Consequently, we have

$$\frac{\Delta Y}{\Delta I} = \frac{1}{1-c} = \text{The multiplier}$$

In words, this says

$$\text{Multiplier} = \frac{1}{1 - \text{Marginal propensity to consume}}$$

or, since

$$\text{Marginal propensity to save} = 1 - \text{Marginal propensity to consume}$$

we have,

$$\text{Multiplier} = \frac{1}{\text{Marginal propensity to save}}$$

A graphical derivation of the multiplier is given in Figure 6–2.

In our numerical example, the marginal propensity to consume is 0.5; consequently the multiplier is 2. Thus, an increase of $10 billion per year in the rate of investment will raise the equilibrium level of income by $20 billion. After this has happened, starting from the initial equilibrium position given above in which investment was $100 billion and income was $440 billion, we will be in the following position (billions of dollars):

Y (= aggregate supply)	460
C (= $0.5Y + 120$)	350
I (new level)	110
S (= $Y - C$)	110
Aggregate demand (= $C + I$)	460

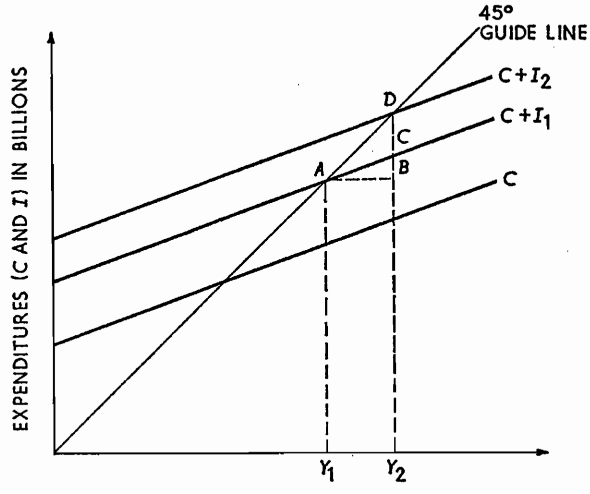
Clearly, this is an equilibrium, since aggregate demand equals aggregate supply, or planned saving equals planned investment.

The multiplier will be larger, the larger is the marginal propensity to consume (smaller is the marginal propensity to save). Thus, the multiplier is 10 if the marginal propensity to consume is 0.9, 5 if the marginal propensity to consume is 0.8, 2 if the marginal propensity to consume is 0.5 (as in our example), and so on.

The multiplier can be explained very easily in commonsense terms. When investment expenditures rise, this will stimulate production and cause income to rise. As income rises, a portion (measured by the marginal propensity to consume) will be spent on consumption, while the remainder (measured by the marginal propensity to save) will be set aside as saving. The consumption spending tends to stimulate production and cause a further rise in income. Equilibrium will be reached when the additional saving caused by the rise in income is sufficient to balance the initial increase in investment, thus restoring the equality between planned saving and planned investment. If (for example) the marginal propensity to save is 20 percent,

FIGURE 6-2

Graphical Derivation of the Multiplier



CD = Change in investment (ΔI)
BC = Change in consumption (ΔC)
$AB = BD$ = Change in income (ΔY)

$$BD = BC + CD$$

$$\text{Marginal propensity to consume} = \frac{BC}{AB} = \frac{BC}{BD}$$

$$BC = BD\frac{BC}{BD}$$

$$BD = BD\frac{BC}{BD} + CD$$

$$BD\left(1 - \frac{BC}{BD}\right) = CD$$

$$BD = \frac{1}{1 - \dfrac{BC}{BD}}CD$$

$$\Delta Y = \frac{1}{1 - \text{Marginal propensity to consume}}\Delta I$$

income will have to rise by $5 in order to induce $1 of additional saving. Consequently, income will have to rise by five times the initial increase in investment in order to induce enough additional saving to restore a balance between planned saving and planned investment. When this has been accomplished, income will stop rising.

The *static* multiplier, to which the present discussion relates, merely tells us how much the *equilibrium* level of income will rise as a result of a sustained increase in investment. Obviously, income does not instantly jump to the new equilibrium when investment increases; rather it is a matter of hiring workers, increasing production, and so on, which takes time. In the next chapter, we will discuss the dynamics of the multiplier, or the nature of the process by which income changes from one level to the other.

INTRODUCTION OF FISCAL VARIABLES

In the above discussion, we deliberately left out government spending and taxing by making the assumption that all output was used either for private consumption or private investment. Let us now make the analysis somewhat more realistic by introducing government activity into our model. We shall begin with a model in which all taxes are lump-sum taxes—that is, the government's tax collections do not change as income changes.[2] With government activity, the equilibrium condition becomes

$$Y = C + I + G \tag{4}$$

where Y is total gross output (or income), C is consumption, I is planned investment and G is government purchases of goods and services. Since all income must be either spent on consumption, saved, or paid to the government in taxes, we have $Y = C + S + T$, where S is planned saving and T is net taxes. Eliminating C between this equation and Equation 5 and transposing terms, we obtain

$$I = S + (T - G) \tag{5}$$

That is, the equilibrium condition can also be stated as "planned investment is equal to planned saving plus the government's surplus (or minus its deficit)." This is an extension of the equilibrium condition "planned investment is equal to planned saving" that applies in an economy without government activity.

Government Purchases of Goods and Services

Since government purchases of goods and services enter into the analysis in exactly the same way as private investment, the multiplier applicable to such government purchases is the same as that applicable to investment as discussed above. Thus, for government purchases, we have

$$\text{Multiplier} = \frac{1}{1 - c}$$

That is, an increase in government spending on, say, public works will stimulate additional production and income, and the rise in income will

[2] An example of a lump-sum tax would be a "head tax" of so much per person. Collections under such a tax would depend only on the size of the population and would not change with income.

continue until it has produced a sufficient rise in planned private saving to balance the rise in government expenditures (reduction of government surplus), so that the equilibrium condition (Equations 4 and 5 above) will be satisfied.

Changes in Taxes and Transfer Payments

A reduction in taxes, like an increase in government purchases of goods and services, will raise the level of aggregate demand and will set off a multiplier effect. Ordinarily, the effects of tax reductions are indirect and somewhat weaker than the effects of increases in expenditures on goods and services. The reason is that a reduction of taxes does not *directly* add anything to aggregate demand, whereas an increase in government purchases of goods and services is itself an increase in aggregate demand which tends directly to increase the national product. A tax reduction increases spending on goods and services only to the extent that it induces the recipients of the tax relief to spend more than they otherwise would have spent.

This can perhaps best be brought out by means of an example. Suppose that the government reduces taxes by $100 million. As a result, taxpayers are left with $100 million more of disposable income than they otherwise would have had. Suppose the marginal propensity to consume is 80 percent. Then taxpayers will increase their consumption expenditures by 80 percent of the increase in their disposable incomes, or by $80 million, while increasing their saving by $20 million. The expenditures of $80 million can be expected to generate $80 million of additional production and income, part of which will be spent on additional consumption, thus initiating the multiplier process. With a marginal propensity to consume of 80 percent, the ordinary multiplier will be 5, and the total expansion will be $400 million (that is, 5 times the initial autonomous increase in consumer spending of $80 million). Consumption will increase by a further $320 million (80 percent of the $400 million increase in income) and saving by a further $80 million. The results may be summarized:

Taxes	−100
Consumption expenditures	400
By those whose taxes were reduced	80
By others	320
Saving	100
By those whose taxes were reduced	20
By others	80

When the process is completed, the government surplus will have been reduced by $100 million (due to the fall in taxes of that amount), while planned saving will have increased by $100 million, so that the condition of equilibrium will be satisfied. In this case, the multiplier for the tax reduction is —4; that is, the decline in taxes of \$100 million has increased the level of income by \$400 million.

This result can easily be generalized. If ΔT is the change in taxes, ΔY is the change in income, and c is the marginal propensity to consume, the change in expenditures resulting directly from the tax adjustment will be $-c\Delta T$. This initial effect will be subject to the ordinary multiplier explained earlier, which will be $1/(1-c)$. Thus, the change in income is given by the following expression:

$$\Delta Y = -\frac{c}{1-c}\Delta T$$

and the multiplier applicable to tax changes is

$$\frac{\Delta Y}{\Delta T} = -\frac{c}{1-c}$$

In the example considered above, this multiplier is $-0.8/(1-0.8)$, or —4.

An increase in transfer payments of \$100 million, with a marginal propensity to consume of 0.8, will likewise raise total income by \$400. In general, the multiplier for transfer payments is $c/(1-c)$, the same as the multiplier applicable to tax changes, except with the opposite sign, because transfer payments are, in essence, negative taxes.

It may be noted that

$$\frac{c}{1-c} = \frac{1}{1-c} - 1$$

That is, the multiplier for transfer payments is one less than the ordinary multiplier applicable to private investment or government purchases of goods and services. Similarly, the multiplier for tax adjustments is one less than the ordinary investment or government purchases multiplier, but with the opposite sign.

The Balanced Budget Multiplier

At first glance, it might seem that if the government were to increase its expenditures while at the same time raising taxes by an equal amount, it would be taking as much out of the expenditure stream through taxes as it was putting in through spending, so that the effect on the level of economic activity would be neutral—i.e., it would neither cause business activity to expand nor cause it to contract. On more careful examination, however, it is apparent that such a balanced increase in expenditures and taxes would usually have an expansive effect.

Suppose, for example, that government purchases of goods and services are increased by \$100 million and that taxes are increased by an equal amount, and suppose further that the marginal propensity to consume is 0.6. Initially, then, government purchases will be increased by \$100 million, while private consumption expenditures will be reduced by \$60 million (60 percent of the \$100 million increase in taxes). Thus, total expenditures will be increased by \$40 million (\$100 million increase in government purchases minus \$60 million decrease in private spending). This net increase in expenditures will be subject to the ordinary multiplier of $1/(1-c)$; with c of 0.6, this multiplier will be 2.5. Thus, the total increase in income will be 2.5 times \$40 million, or \$100 million. That is, an increase of \$100 million in both expenditures and taxes will raise equilibrium income by \$100 million, or the *balanced budget multiplier* will be unity.

In the case in which taxes are of the lump-sum variety, not affected by changes in income, it is easy to show that the balanced budget multiplier is equal to unity regardless of the value of c, the marginal propensity to consume. That is, when we add the multiplier for government purchases, $1/(1-c)$, and the multiplier for taxes, $-c/(1-c)$, we have

$$\frac{\Delta Y}{\Delta B} = \frac{1}{1-c} - \frac{c}{1-c} = \frac{1-c}{1-c} = 1$$

where ΔB represents an equal increase in government purchases and in taxes. The student may verify for himself that in such a case it would be necessary to raise taxes by $1/c$ times an increase in government purchases in order to prevent any expansion of aggregate demand. The analysis applies, of course, in the opposite direction, that is, a balanced reduction in government purchases and in taxes will lower the equilibrium level of income by the amount of the reduction.

Taxes Responsive to Income

In the above discussion, we have made the unrealistic assumption that all the taxes and transfer payments are of the lump-sum variety, or at least of some kind not dependent on income. In practice, nearly all of the taxes which are used are in one degree or another responsive to income—that is, tax collections rise with income. Moreover, some types of transfer payments, such as unemployment compensation benefits, tend to fall as the level of income and employment rises and to rise as the level of income and employment falls. We shall now adapt our analysis in a rather simple way to allow for this dependence of tax collections and transfer payments on the level of income.

To begin with, the change in total income in any period is equal to the sum of the change in consumption expenditures (ΔC) plus the change in investment expenditures (ΔI) plus the change in government purchases of goods and services (ΔG). That is,

$$\Delta Y = \Delta C + \Delta I + \Delta G \tag{6}$$

Since we assume that consumption is dependent on disposable personal income (Y_d), we have

$$\Delta C = c\,\Delta Y_d \tag{7}$$

For the present, we shall make the oversimplified assumption that to derive disposable income from GNP (Y) it is only necessary to subtract taxes (X) and add government transfer payments (F). That is,

$$\Delta Y_d = \Delta Y - \Delta X + \Delta F$$

Let us define net taxes (T) as taxes minus transfer payments, i.e., treat transfer payments as negative taxes. Thus,

$$\Delta T = \Delta X - \Delta F$$

or

$$\Delta Y_d = \Delta Y - \Delta T \tag{8}$$

If we treat taxes and transfer payments as partly dependent upon GNP and partly autonomous and therefore subject to adjustment by government, we have

$$\Delta T = t\,\Delta Y + \Delta T_a \tag{9}$$

where ΔT is the total change in net taxes, ΔT_a is the autonomous element in net taxes, and t may be called the "the government's net marginal propensity to collect taxes." The size of t will, of course, depend upon the nature of the laws governing taxes and transfer payments. For example, if a progressive personal income tax, which is highly sensitive to changes in income, is an important part of the tax system, this will tend to make t large.[3] Similarly, a liberal unemployment compensation system, which pays large benefits to workers when they become unemployed, will make t large by causing transfer payments to rise sharply (thus causing *net* taxes to fall) when income falls.

Substituting from Equations 7, 8, and 9 into Equation 6, we obtain

$$\Delta Y = c\,\Delta Y - ct\,\Delta Y - c\,\Delta T_a + \Delta I + \Delta G$$

or

$$\Delta Y(1 - c + ct) = -c\,\Delta T_a + \Delta I + \Delta G$$

[3] A *progressive tax* is one for which the ratio of tax payments to income increases as income increases; that is, taxes change more than in proportion to income. A *regressive tax* is one under which the ratio of tax payments to income falls as income rises, while a *proportional tax* is one under which the ratio of tax payments to income remains constant as income changes.

or

$$\Delta Y = \frac{1}{1 - c(1 - t)} [-c\Delta T_a + \Delta I + \Delta G]$$

From this, it is apparent that the multiplier applicable to a *change in private investment* (ΔI) or a *change in government purchases of goods and services* (ΔG) is

$$\frac{\Delta Y}{\Delta I} = \frac{\Delta Y}{\Delta G} = \frac{1}{1 - c(1 - t)}$$

while the multiplier applicable to *autonomous changes in the level of net taxes* (ΔT_a) is

$$\frac{\Delta Y}{\Delta T_a} = \frac{-c}{1 - c(1 - t)}$$

The values of the multiplier applicable to autonomous expenditures—private investment or government purchases of goods and services—for various

TABLE 6–1

Size of Multiplier Applicable to Autonomous Expenditures for Various Values of c and t

Marginal Propensity to Consume (c)	*Government's Marginal Net Propensity to Collect Taxes* (t)					
	0	0.1	0.2	0.33	0.4	0.5
0.4	1.67	1.56	1.47	1.36	1.32	1.25
0.5	2.00	1.82	1.67	1.50	1.43	1.33
0.6	2.50	2.17	1.92	1.67	1.56	1.43
0.7	3.33	2.70	2.27	1.88	1.72	1.54
0.8	5.00	3.57	2.78	2.14	1.92	1.67
0.9	10.00	5.26	3.57	2.50	2.17	1.82

values of the marginal propensity to consume (c) and of the government's marginal net propensity to collect taxes (t) are given in Table 6–1. As indicated, the more sensitive are taxes and transfer payments to changes in the level of income, i.e., the larger is t, the smaller is the multiplier. Thus, the larger t is, the less sensitive is the level of income to autonomous changes in investment and government expenditures.

This effect of taxes in reducing the size of the multiplier is easy to understand once it is recognized that $c(1 - t)$ is the marginal propensity to consume *out of GNP*. Thus, if the government's marginal propensity to collect taxes is 20 percent, an increase of \$1 billion in GNP will cause tax collections to rise by \$200 million, leaving an increase of only \$800 million in disposable personal income. If the marginal propensity to consume out of disposable income is 80 percent, personal consumption expenditures will increase by \$640 million dollars or 64 percent of GNP. By the standard formula,

$$\text{Multiplier} = \frac{1}{1 - \text{Marginal propensity to consume}}$$

the multiplier works out to be $1/(1 - 0.64) = 2.78$, the value shown in Table 6–1 for $c = 0.8$ and $t = 0.2$.

A More Complete Model

The above analysis is greatly oversimplified in allowing, in effect, for only a single type of tax which is assumed to depend on GNP, and in lumping

TABLE 6–2

Relation between GNP and Disposable Personal Income

Symbol Used	*National Income Accounts Category*	*1968 Amount (Billions)*
Y	Gross National Product	\$865.7
A	Less: Capital consumption allowances	73.3
T_i	Indirect business taxes [a]	77.1
P	Corporate profits and inventory valuation adjustment	87.9
T_s	Contributions for social insurance	47.0
T_p	Personal taxes	97.9
—	Statistical discrepancy	−2.5
F_p	Plus: Government transfer payments to persons	55.8
$G_i + C_i$	Interest paid by government (net) and by consumers	26.1
D	Dividends	23.1
Y_d	Equals: Disposable Personal Income	\$590.0

[a] In the interest of simplicity, subsidies less current surplus of government enterprises (\$0.8 billion) has been treated as negative indirect business taxes—i.e., it has been deducted from indirect business taxes (\$77.9 billion) to obtain the amount shown.

together personal and corporate saving. It is possible to improve substantially on this model without introducing too many complexities into the analysis.

Table 6–2 shows the adjustments that must be made to go from GNP to personal disposable income.[4] In terms of the symbols shown in the left-hand column of the table, the relation can be written in the form of an equation as follows (neglecting the statistical discrepancy):

[4] See the discussion in Chapter 2, pp. 31–36, and in Chapter 3, pp. 42–49, and Table 3–1, p. 55.

$$Y_d = Y - A - T_i - P - T_s - T_p + F_p + G_i + C_i + D$$

It will simplify the development of the multiplier considerably to express this equation in terms of changes; that is

$$\Delta Y_d = \Delta Y - \Delta A - \Delta T_i - \Delta P - \Delta T_s - \Delta T_p + \Delta F_p + \Delta G_i + \Delta C_i + \Delta D \quad (10)$$

That is, the change in disposable personal income during any period of time is necessarily equal to the algebraic sum of the changes in GNP and the other variables shown on the right-hand side of this equation.

Capital consumption allowances (A), which consist mainly of depreciation charges, depend primarily on the size of the capital stock. Since the capital stock will be affected only indirectly and quite slowly by the forces we are interested in, it will be justifiable to treat capital consumption allowances as a constant. Similarly, we will treat government and consumer interest payments (G_i and C_i) as constants, on the ground that they are little affected by the changes we are analyzing. Thus, we assume that $\Delta A = \Delta G_i = \Delta C_i = 0$, and Equation 10 reduces to

$$\Delta Y_d = \Delta Y - \Delta T_i - \Delta P - \Delta T_s - \Delta T_p + \Delta F_p + \Delta D \quad (11)$$

Indirect taxes (T_i) consists mainly of excise taxes and taxes on business property. Property taxes are not affected directly by changes in income, while excise taxes depend on sales of the taxed goods and services. Let us suppose that sales of items subject to excise taxes depend on GNP (Y). This gives us an equation

$$\Delta T_i = t_i \Delta Y \quad (12)$$

where t_i is a coefficient measuring the marginal response of indirect business taxes to changes in GNP.

Corporate profits (P) respond positively to changes in total production and income; that is, when GNP changes, corporate profits change by some fraction of the change in GNP. Thus, we have

$$\Delta P = p \Delta Y \quad (13)$$

where p is the marginal response of profits to changes in GNP.

Contributions for social insurance (T_s) are in essence taxes on payrolls. Since employment, and therefore payrolls, rise and fall with GNP, we can express the relation as

$$\Delta T_s = t_s \Delta Y \quad (14)$$

where t_s is the marginal response of contributions for social insurance to changes in GNP.

Personal taxes (T_p) consist primarily of income taxes which are levied at the federal and state levels. The behavior of such taxes depends upon the way in which aggregate taxable income changes, on the way in which the distribution of taxable income among incomes in various tax brackets changes, and on the set of tax rates applicable to income in various brackets. Since excise and property taxes are generally deductible in computing taxable income, while social insurance contributions are not deductible and government transfer payments are not subject to tax, we may define taxable income (Y_t) for purposes of individual income taxes, as follows:

$$Y_t = Y - T_i - P + D$$

or, in terms of changes,

$$\Delta Y_t = \Delta Y - \Delta T_i - \Delta P + \Delta D \tag{15}$$

Then we have

$$\Delta T_p = t_p \Delta Y_t + \Delta T_{pa} \tag{16}$$

where t_p is the government's marginal propensity to collect additional personal taxes out of taxable income, and ΔT_{pa} is an autonomous element that is introduced to handle changes in the level of personal taxes.

Some forms of personal transfer payments (F_p), particularly unemployment compensation benefits, respond inversely to changes in GNP. That is, a rise in GNP will bring an increase in employment and a reduction in unemployment. As unemployed workers find jobs, they will cease to draw unemployment compensation benefits. Conversely, when GNP declines, unemployment will increase and so will unemployment compensation benefits. Thus, we have

$$\Delta F_p = f_p \Delta Y \tag{17}$$

where f_p is the marginal response of personal transfer payments to changes in GNP. Note that f_p is negative, that is, benefit payments decline as GNP increases.

Corporate dividend payments are assumed to depend upon the amount of corporate profits remaining after payment of corporate profits taxes. Thus,

$$\Delta D = d \Delta P_{at} \tag{18}$$

where D is dividends and P_{at} is corporate profits after tax. The coefficient d is the marginal propensity of corporations to pay dividends out of aftertax profits. Corporate aftertax profits are defined to be equal to corporate profits before tax (P) minus corporate profits taxes (T_c). That is,

$$\Delta P_{at} = \Delta P - \Delta T_c \tag{19}$$

Corporate profits taxes, in turn, depend on corporate profits before tax; thus

$$\Delta T_c = t_c \Delta P + \Delta T_{ca} \tag{20}$$

where t_c is the marginal tax rate applicable to corporate profits, and ΔT_{ca} is an autonomous element that is introduced to enable us to analyze the effects of changes in the level of corporate profits taxes.

Finally, we have a consumption function and an equation expressing GNP as the sum of consumption, investment, and government purchases of goods and services. Expressed in terms of differences, these equations are

$$\Delta C = c\,\Delta Y_d \tag{21}$$

$$\Delta Y = \Delta C + \Delta I + \Delta G \tag{22}$$

where C is personal consumption expenditures, I is gross private domestic investment, G is government purchases of goods and services, and c is the marginal propensity to consume out of disposable personal income. I and G are exogenous variables; that is, they are determined outside of the model. We can therefore use the model to analyze the effects on GNP of autonomous changes in I or G.

By solving this system of 12 equations (Equations 11–22), it is possible to derive multipliers which show how GNP will be affected by autonomous changes in investment (I), government purchases (G), the level of personal taxes (T_{pa}), and the level of corporate income taxes (T_{ca}).

The procedure for solving the equations is as follows: First, substitute Equations 12, 13, 14, 16, 17, and 18 into Equations 11 and 15. This process yields

$$\Delta Y_d = \Delta Y - t_i\,\Delta Y - p\,\Delta Y - t_s\,\Delta Y - t_p\,\Delta Y_t - \Delta T_{pa} + f_p\,\Delta Y + d\,\Delta P_{at} \tag{23}$$

$$\Delta Y_t = \Delta Y - t_i\,\Delta Y - p\,\Delta Y + d\,\Delta P_{at} \tag{24}$$

From Equations 13, 19, and 20, we have

$$\Delta P_{at} = p(1 - t_c)\,\Delta Y - \Delta T_{ca} \tag{25}$$

Substituting Equation 25 into Equations 23 and 24 and then substituting Equation 24 into Equation 23, we obtain, after some simplification,

$$\Delta Y_d = [(1 - t_p)(1 - t_i - p(1 - d(1 - t_c))) - t_s + f_p]\,\Delta Y - \Delta T_{pa} - d(1 - t_p)\,\Delta T_{ca} \tag{26}$$

Substituting Equation 26 into Equation 21, then substituting Equation 21 into Equation 22 and solving for ΔY, we obtain

$$\Delta Y = \frac{1}{1 - c[(1 - t_p)(1 - t_i - p(1 - d(1 - t_c))) - t_s + f_p]}\,[-c\Delta T_{pa} - cd(1 - t_p)\Delta T_{ca} + \Delta I + \Delta G] \tag{27}$$

From this expression, we obtain the following four multipliers:

1. *For an autonomous change in investment* (ΔI): setting $\Delta T_{pa} = \Delta T_{ca} = \Delta G = 0$ in Equation 27, we obtain

$$\Delta Y = \frac{1}{1 - c[(1 - t_p)(1 - t_i - p(1 - d(1 - t_c))) - t_s + f_p]} \Delta I \quad (28)$$

2. *For a change in government purchases* (ΔG): setting $\Delta T_{pa} = \Delta T_{ca} = \Delta I = 0$ in Equation 27, we have

$$\Delta Y = \frac{1}{1 - c[(1 - t_p)(1 - t_i - p(1 - d(1 - t_c))) - t_s + f_p]} \Delta G \quad (29)$$

3. *For a change in the level of personal income taxes* (ΔT_{pa}): setting $\Delta T_{ca} = \Delta I = \Delta G = 0$ in Equation 27, we have

$$\Delta Y = \frac{-c}{1 - c[(1 - t_p)(1 - t_i - p(1 - d(1 - t_c))) - t_s + f_p]} \Delta T_{pa} \quad (30)$$

4. *For a change in the level of corporate income taxes* (ΔT_{ca}): setting $\Delta T_{pa} = \Delta I = \Delta G = 0$ in Equation 27, we obtain

$$\Delta Y = \frac{-cd(1 - t_p)}{1 - c[(1 - t_p)(1 - t_i - p(1 - d(1 - t_c))) - t_s + f_p]} \Delta T_{ca} \quad (31)$$

To give concreteness to the discussion, let us consider a numerical example. The values of the parameters are intended to approximate those prevailing in the United States in a period such as 1961–65 when the economy was expanding but there existed some margin of unemployed resources. They are: $c = 0.90$, $t_p = 0.16$, $t_i = 0.08$, $p = 0.14$, $d = 0.40$, $t_c = 0.42$, $t_s = 0.02$, and $f_p = -0.02$. Taking, for example, a change in government purchases as the initiating force, we have from Equation 26, after setting $\Delta T_{pa} = \Delta T_{ca} = 0$, since we are assuming no change in the tax functions,

$$\Delta Y_d = [(1 - t_p)(1 - t_i - p(1 - d(1 - t_c))) - t_s + f_p]\, \Delta Y \quad (32)$$

Substituting the assumed values of the parameters into this equation, we obtain

$$\Delta Y_d = 0.6425\, \Delta Y \quad (33)$$

That is, each change of $1 billion in GNP will cause a change of only $642.5 million in the same direction in disposable personal income. The reason for this is that portions of the increase in income generated by the rise in GNP are drained off in the form of various kinds of taxes paid to governments (federal, state, and local); transfer payments decline as unemployment is reduced; and corporations retain a portion of their additional aftertax profits instead of paying the entire amount to their stockholders as dividends. Table 6–3, which is organized in the same way as Table 6–2, shows how all of the items lying between GNP and disposable personal income are affected by an assumed $1 billion increase in GNP. The student

TABLE 6-3

Calculation of Effects on Disposable Personal Income Resulting from a $1 Billion Increase in GNP

Symbol Used	*National Income Category*	*How Calculated*	*Amount of Change (Billions)*
ΔY	Gross National Product	Assumed	$1.0000
ΔA	Less: Capital consumption allowances	...	...
ΔT_i	Indirect business taxes	$t_i\Delta Y$	0.0800
ΔP	Corporate profits	$p\Delta Y$	0.1400
ΔT_s	Social insurance contributions	$t_s\Delta Y$	0.0200
ΔT_p	Personal taxes	Note (1)	0.1300
ΔF_p	Plus: Government transfer payments to persons	$f_p\Delta Y$	−0.0200
$\Delta G_i + \Delta C_i$	Interest paid by government and consumers	...	...
ΔD	Dividends	Note (2)	0.0325
ΔY_d	Equals: Disposable Personal Income		$0.6425

Assumed values of parameters: $t_i = 0.08$, $p = 0.14$, $d = 0.40$, $t_c = 0.42$, $f_p = -0.02$, $t_p = 0.16$.

NOTE (1): Calculation of change in personal taxes:

$$\Delta Y_t = \Delta Y - \Delta T_i - \Delta P + \Delta D$$
$$\Delta Y_t = 1.0000 - 0.0800 - 0.1400 + 0.0325 = 0.8125$$
$$\Delta T_p = t_p\Delta Y_t = 0.1300$$

NOTE (2): Calculation of change in dividends:

$$\Delta T_c = t_c\Delta P = 0.0588$$
$$\Delta P_{at} = \Delta P - \Delta T_c = 0.1400 - 0.0588 = 0.0812$$
$$\Delta D = d\Delta P_{at} = 0.0325$$

is urged to work through the calculations underlying this table, since this should help him to understand the analysis.

As shown by Equation 28 above, the multiplier for an autonomous change in private investment is

$$\frac{\Delta Y}{\Delta I} = \frac{1}{1 - c[(1 - t_p)(1 - t_i - p(1 - d(1 - t_c))) - t_s + f_p]}$$

As already shown, the value of the expression in brackets in the denominator is 0.6425. Since the marginal propensity to consume out of disposable personal income c is assumed to have a value of 0.90, the multiplier becomes

$$\frac{\Delta Y}{\Delta I} = \frac{1}{1 - (0.90)(0.6425)} = \frac{1}{0.42175} = 2.37 \tag{34}$$

That is, the multiplier for a change in private investment in this model is approximately 2.4.

This result can be related rather easily to the simpler model developed earlier in this chapter. We saw (page 118 above) that the simple multiplier is equal to

$$\frac{1}{1 - \text{Marginal propensity to consume}}$$

In order for this simple formula to hold, the marginal propensity to consume *out of GNP,* that is, $\Delta C/\Delta Y$, must be used. By simple algebra, we have

$$\frac{\Delta C}{\Delta Y} = \frac{\Delta C}{\Delta Y_d}\frac{\Delta Y_d}{\Delta Y}$$

That is, the marginal propensity to consume out of GNP is equal to the marginal propensity to consume out of disposable personal income multiplied by the increase in disposable personal income produced by each dollar rise in GNP. In our numerical example, $\Delta C/\Delta Y_d = 0.90$, and we have shown in Equation 33 that $\Delta Y_d/Y = 0.6425$. Thus, we have

$$\frac{\Delta C}{\Delta Y} = (0.90)(0.6425) = 0.57825$$

That is, in this case, given the tax, transfer payment, corporate profit, and dividend responses, the effective marginal propensity to consume out of GNP is 0.57825. Substitution of this value into the simple multiplier formula yields the value 2.37.

Multipliers for Government Purchases and Tax Changes

The multiplier for a change in government purchases of goods and services is given by Equation 29 and is the same as the multiplier applicable to a change in investment. Equations 30 and 31 give the multipliers for changes in the level of personal income taxes and corporate profits taxes. The multiplier for a change in personal income taxes (Equation 30) has a negative sign—that is, a reduction in taxes increases income—and is equal in absolute value to the multiplier for government purchases multiplied by the marginal propensity to consume. Thus, in our example, the marginal propensity to consume is 0.90, and the multiplier for a change in personal taxes is therefore

$$\frac{\Delta Y}{\Delta T_{pa}} = -(0.90)(2.37) = -2.13$$

The multiplier for a change in the level of corporate profits taxes (Equation 31) also has a negative sign and is equal in absolute value to the multiplier for government purchases multiplied by $c(1 - t_p)d$. In our illustration, this has a value of $-(0.90)(0.40)(1 - 0.16) = -0.3024$. The multiplier for a change in the level of corporate taxes is therefore

$$\frac{\Delta Y}{\Delta T_{ca}} = -(0.3024)(2.37) = -0.717$$

That is, a reduction of $1 billion in corporate profits taxes would raise GNP by about $717 million. The expression $c(1 - t_p)d$ can be understood quite easily. If corporate profits taxes are cut by, say, $1 billion, this will increase aftertax profits by $1 billion, and corporations are assumed to pay out d percent of this $1 billion in additional dividends. Personal taxes will absorb t_p percent of the additional dividends, leaving $1 - t_p$ percent as an addition to disposable personal income, and c percent of the increase in disposable personal income will be spent for added consumption. Thus, the initial impact of a $1 billion tax cut on consumption will be $c(1 - t_p)d$ billion dollars—in our example, $302.4 million. The basic multiplier (given by Equation 28 or Equation 29) of 2.37 will apply to this initial increase in spending.

While the multiplier applicable to a change in corporate profits taxes is relatively small, it should be especially noted in this connection that the analysis we have developed deals only with effects on consumption resulting from induced changes in dividends. A change in corporate profits taxes, by changing aftertax profits and retained corporate earnings, may also have a substantial effect on investment which we are neglecting.

In practice, most changes in personal or corporate profits taxes that would be made would change not only the level of the relevant tax function (T_{pa} or T_{ca}) but also the coefficients of response of taxes to income (t_p or t_o). We shall not, however, attempt to deal with this problem. As a practical matter, many changes in taxes will not alter t_p or t_c very greatly, and multipliers of the type developed above are therefore often fairly satisfactory in dealing with the effects of tax changes.

Finally, the balanced budget multiplier for equal changes in government purchases and in the level of personal income taxes is worth mentioning. This multiplier is equal to the sum of Equations 29 and 30, which is

$$\frac{\Delta Y}{\Delta B} = \frac{1 - c}{1 - c[(1 - t_p)(1 - t_i - p(1 - d(1 - t_c))) - t_s + f_p]}$$

where ΔB stands for an equal change in government purchases and in the level of personal income taxes. In our example the value of $1 - c$ is 0.10, and this multiplier works out to be

$$\frac{\Delta Y}{\Delta B} = (0.10)(2.37) = 0.237$$

That is, a balanced increase in government purchases and in personal taxes of $1 billion would raise GNP by only about $237 million. Thus, while in the unrealistic case of lump-sum taxes discussed earlier, the balanced budget multiplier is unity, in more realistic cases it turns out to be too small to be of much importance.

Automatic Stabilizers

With a fiscal system of the kind we have, a rise in income automatically causes the government budget to shift toward a surplus as tax revenues increase and certain types of transfer payments, particularly unemployment compensation benefits, decline. The rise in taxes and the fall in transfer payments drain purchasing power away from households, with the result that disposable personal income rises less than GNP. By limiting the movements of disposable personal income, the fiscal system causes personal consumption expenditures to change less than would be the case in the absence of these tax and transfer payment responses, thereby reducing the size of the multiplier. Taxes and transfer payments that function in this way are called *automatic fiscal stabilizers* since they reduce the size of the fluctuations in GNP resulting from autonomous fluctuations in private investment or consumption demand. In addition, the tendency of corporations to pursue a stable dividend policy—that is, to change dividend payments in either direction by less than the change in aftertax profits—introduces an element of automatic stabilization into the private sector of the economy. In our numerical example, it is the presence of these automatic stabilizers in the public and private sectors that causes disposable personal income to change by only $0.6425 for each $1 change in GNP, as shown in Table 6–3.

The student can verify from Equation 28 that if $t_p = t_i = t_c = t_s = f_p = 0$ and $d = 1$, the multiplier applicable to a change in private investment would be

$$\frac{\Delta Y}{\Delta I} = \frac{1}{1 - c}$$

That is, if all taxes and personal transfer payments were of the lump-sum variety, not responsive, directly or indirectly, to changes in GNP, and if corporations followed the practice of adjusting their dividend payments by the full amount of any change in their profits, the multiplier would take on the simple form developed earlier in this chapter. With the assumed marginal propensity to consume out of disposable personal income of 90 percent, the multiplier would be 10. (Note that the multiplier in this case is not affected by p, the responsiveness of profits to changes in GNP, since any increment to profits is entirely paid out in dividends and therefore becomes a part of disposable personal income.)

Thus, it is apparent that the presence of the various taxes and transfer payments at the rates assumed and of the specified corporate policy with respect to payment of dividends cuts the multiplier from 10 to 2.37, a reduction of 76.3 percent. That is, the change in income produced by an autonomous change in investment is about 76 percent smaller than it would be with lump-sum taxes and transfer payments and with a corporate policy of paying all increases in profits to stockholders.

There are some additional factors, not taken into account in the above discussion, which help to stabilize the economy in the face of cyclical fluctuations stemming from variations in private investment. In calculating the multipliers presented above, we have chosen values for the response coefficients, c, p, and d, that are roughly characteristic of periods when the economy is expanding under conditions fairly close to full employment. Thus, the multipliers we have presented would be reasonably suitable as applied to an increase in government expenditures, a reduction in taxes, or an autonomous increase in investment in a period such as 1961–65 when the economy was already expanding but still had some margin of unutilized resources. If we were considering the effects of a sharp fall in investment sufficient to push the economy into a recession, there would be several adjustments that should be made in our calculations. In the first place, we have employed a marginal propensity to consume out of disposable personal income of 90 percent, which is reasonably appropriate under conditions in which the economy is expanding with a relatively small margin of unemployed resources; however, as explained in Chapter 5, in a recession the marginal propensity to consume would be likely to fall, perhaps to approximately 70 percent. Second, while the responses of profits and dividends employed above are reasonably appropriate to a period of expansion, these responses are likely to be quite different in the course of a recession. Profits typically respond very sensitively during cyclical fluctuations, and corporations have a marked tendency to maintain dividends during recessions. Thus, during a recession p is likely to be substantially greater and d smaller than the values of 0.14 and 0.40 used above. Finally, in a recession, unemployment will increase sharply and the increase in unemployment benefits will be substantially more than 2 percent of the fall in GNP, as assumed above.

A tendency for a large portion of a change in GNP to be reflected in a change in corporate profits means that other forms of income will be rendered more stable. Changes in corporate profits will affect disposable personal income only to the extent that they are reflected in changes in dividends, whereas changes in other types of incomes will be more directly reflected in disposable personal income. Consequently, the larger the response of corporate profits to changes in GNP, the greater will be the stability of disposable personal income if other responses are taken as given. Moreover, since dividends enter directly into disposable personal income, a small response of dividends to changes in corporate profits after taxes will also help to stabilize disposable personal income, again assuming other responses are given. Thus a large value of p, the marginal response of profits to GNP, and a small value of d, the marginal propensity of corporations to pay dividends out of aftertax profits, will help to reduce the change in disposable personal income produced by a given change in GNP. A large (absolute) value of f_p, the response of personal transfer payments to GNP, will have a similar effect. It will be useful to repeat Equation 32, which shows the relation between GNP (Y) and disposable personal income (Y_d):

$$\Delta Y_d = [(1 - t_p)(1 - t_i - p(1 - d(1 - t_c))) - t_s + f_p]\,\Delta Y \quad (32)$$

(The student should satisfy himself that an increase in p, a reduction in d, or an increase in the absolute size of f_p will reduce the size of the expression enclosed in brackets.) Suppose profits change by 35 percent of the change in GNP, dividends change by only 10 percent of the change in aftertax corporate profits, and unemployment compensation benefits increase by 10 percent of the decline in GNP. That is, suppose $p = 0.35$, $d = 0.10$, and $f_p = -0.10$—values that are quite reasonable during a recession period. Assuming the same values for the other parameters that were used above ($t_p = 0.16$, $t_i = 0.08$, $t_c = 0.42$, and $t_s = 0.02$), and substituting in Equation 32, we obtain

$$\Delta Y_d = 0.3759\,\Delta Y$$

That is, when p is increased to 0.35 from the value of 0.16 assumed earlier, d is reduced from 0.40 to 0.10, and f_p is changed from —0.02 to —0.10, the change in disposable income is reduced from 64.25 percent of the change in GNP (see Equation 37) to 37.59 percent—a marked increase in the degree of stabilization.

Let us assume further that the marginal propensity to consume has a value of 70 percent, which would be more appropriate to recession than the 90 percent assumed earlier. By the same reasoning used earlier (see Equation 34), the multiplier would be

$$\frac{\Delta Y}{\Delta I} = \frac{1}{1 - (0.70)(0.3759)} = \frac{1}{0.7369} = 1.36$$

The student is warned not to take the calculations given above as anything more than rough approximations to reality. Although the model may seem rather complex, a fully satisfactory analysis requires an even more sophisticated approach.[5] However, it is quite clear that the multiplier during

[5] For such an analysis, see J. S. Duesenberry, Otto Eckstein, and Gary Fromm, "A Simulation of the United States Economy in Recession," *Econometrica,* Vol. XXVIII, October 1960, pp. 749–809, reprinted in condensed form in R. A. Gordon and L. R. Klein (eds.), *Readings in Business Cycles* (Homewood, Ill.: Richard D. Irwin, Inc., 1965), pp. 237–77. This paper provides an estimate of 1.34 for the multiplier during the course of the business cycle (pp. 761–63 in original; pp. 245–46 in *Readings*), a result very similar to that given above. It also contains an excellent discussion of the automatic stabilizers. Another very useful analysis of the automatic stabilizers is to be found in Wilfred Lewis, Jr., *Federal Fiscal Policy in Postwar Recessions* (Washington: Brookings Institution, 1962), chaps. ii and iii.

a period of recession and recovery is smaller than the multiplier that applies during a period of high-level expansion, and the estimates given above are probably of the right general order of magnitude.[6]

[6] One further factor, not considered above, that is likely to increase the economy's resistance to recessions is the persistent growth of population. It was shown in Figure 5–2 that a rise in population is likely to raise the consumption function, that is, to increase the level of personal consumption expenditures associated with any given level of disposable personal income. Since population can be depended upon to go on growing during a recession, this force will serve to sustain consumption, thereby helping to check the decline in aggregate demand and shorten the recession. (For a further discussion of this point, see Duesenberry, Eckstein, and Fromm, *op. cit.*, p. 771 in original, pp. 251–52 in *Readings.*) Note, however, that this force is asymmetrical; it will be a *destabilizer* during periods of expansion since it will go on raising consumption expenditures under these conditions also.

Chapter 7 THE SHORT-RUN DYNAMICS OF INCOME CHANGE

There is nothing automatic or magic about the multiplier. The expansion of income and employment involves the making of decisions by a multitude of economic units to increase production rates and to hire additional workers, and the completion of time-consuming activities and processes required to produce additional goods and services. It is, therefore, useful to look at the *multiplier meachanism as a dynamic process,* using a motion picture camera rather than merely taking snapshots of successive equilibrium positions.[1]

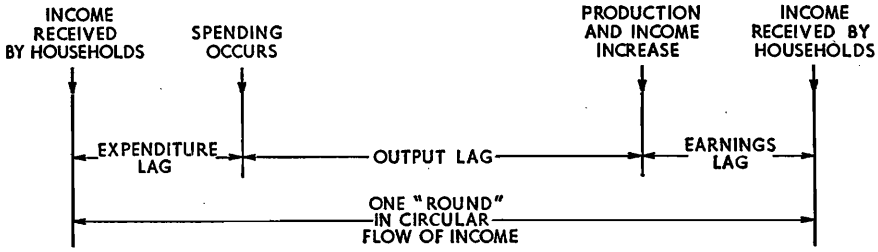
Lags in the Circular Flow of Income

The multiplier may be visualized in terms of successive "rounds" of income flow. Of the income received by households at a particular point in time, a portion is spent on consumption. This expenditure will, in the normal course of events, stimulate the production of additional goods and services, and this production will generate an equivalent amount of income in the form of wages, rent, interest, and profits. A portion of this income will be paid to the government in taxes, another portion will be held back by the corporate business sector as retained earnings, and the remainder will, in due course, be received by households. We may think of a "round" in the process of income generation as the average time that elapses between one household receipt of income and the next in this circular flow.

It is possible to distinguish three lags in this circular flow process. First there is an average lag between the time income is received by housholds and the time a portion of it is spent, which may be called the *expenditure lag.* Second, there is an average lag between the time the spending occurs and the time the production stimulated by this spending takes place and the resulting new incomes are earned, which may be termed the *output lag.* Third, there is an average lag between the time income is earned and the

[1] For a fuller discussion of many of the issues taken up here, see R. M. Goodwin, "The Multiplier," in S. E. Harris (ed.), *The New Economics* (New York: Alfred A. Knopf, 1947), pp. 482–99.

time the requisite portion of it is actually received by households, which we may call the *earnings lag*.[2] The three lags are illustrated conceptually in the following diagram:



Such evidence as there is suggests that the expenditure lag and the earnings lag are not long enough to be very significant.[3] That is, to the extent that consumers buy goods and services out of income, it appears that the time elapsing between the receipt of income and its expenditure is quite short; moreover, for most forms of income, the average time span between the earning of the income and its actual receipt is negligible. Accordingly, we shall disregard the expenditure lag and the earnings lag and base our analysis entirely on the output lag, which appears to be of considerable importance.

A Model with Output Lag

We shall not attempt to work out the dynamics of the rather complex model developed at the end of Chapter 6. Instead, we shall use the less elaborate model employed earlier in that chapter which includes autonomous government expenditures and a single tax that is responsive to income. Later on, we shall show, however, that the same principles can be applied to our more elaborate model.

Since we are assuming that there is no expenditure lag, we shall postulate that consumption (C) in a particular period (t) depends upon disposable personal income (Y^d) in the same period; that is,

$$C_t = cY_t^d + C_0$$

Disposable personal income in period t is equal to gross national product (Y) in that period minus net taxes (T) paid in that period; that is,

$$Y_t^d = Y_t - T_t$$

[2] L. A. Metzler, "Three Lags in the Circular Flow of Income," in *Income, Employment, and Public Policy: Essays in Honor of Alvin H. Hansen* (New York: W. W. Norton & Co., 1948), pp. 11–32.

[3] *Ibid.*

We shall suppose that net tax payments in period t are a linear function of income earned—and also received, since we are assuming the earnings lag is zero—during that period, according to the following relationship:[4]

$$T_t = xY_t + T_0$$

We are assuming that there is an output lag in the model. That is, businessmen adjust their production to their sales, i.e., the expenditures on their products, with a lag of one period. Total production (and total gross income) is designated as Y, and total sales include sales to consumers (C) plus sales to business for investment (I) plus sales to government (G). Thus, since production in a given period is equal to sales in the pervious period, we have

$$Y_t = C_{t-1} + I_{t-1} + G_{t-1}$$

The best way to illustrate the working of such a model is to insert hypothetical values of the marginal propensity to consume (c) and the marginal propensity to pay taxes (x). Let us suppose that $c = 0.8$ and $x = 0.25$, and that $C_0 = 0$ and $T_0 = -40$, so that the model becomes

$$\begin{aligned} C_t &= 0.8Y_t^d \\ Y_t^d &= Y_t - T_t \\ T_t &= 0.25Y_t - 40 \\ Y_t &= C_{t-1} + I_{t-1} + G_{t-1} \end{aligned}$$

In static equilibrium, the values of all the variables—Y, Y^d, C, and T—will remain constant; that is, they will be the same in period t as in period $t - 1$. Accordingly, the static counterpart of the above model assuming I and G are constant at 120 and 168 respectively, is obtained by dropping the time subscripts, which gives us the following:

$$\begin{aligned} C &= 0.8Y^d \\ Y^d &= Y - T \\ T &= 0.25Y - 40 \\ Y &= C + 120 + 168 \end{aligned}$$

This can be solved for the equilibrium value of Y by substituting from the first three equations into the fourth; upon doing this we obtain

$$\begin{aligned} Y &= 0.8(Y - 0.25Y + 40) + 120 + 168 \\ Y(1 - 0.8 + (0.8)(0.25)) &= 32 + 120 + 168 \\ 0.4Y &= 320 \\ Y &= 800 \end{aligned}$$

The equilibrium position is shown in the first two columns of Table 7–1; the reader can see for himself that this is an equilibrium position by

[4] The student may verify for himself that the tax transfer system is *progressive* (see footnote 3, Chapter 6) if T_0 is negative, *regressive* if T_0 is positive, and *proportional* if T_0 is zero.

checking through the calculations to show that the values in the second period are the same as those in the first. That is, when this position is attained, it tends to generate itself over and over again until some autonomous change interrupts the process.

In the third period, such an autonomous change is introduced—investment (I) rises by $10 billion per period—from $120 billion to $130 billion. Thus, total sales rise to $810 billion in that period, and accordingly total output and income rise to $810 billion in the following period, as businessmen adjust their production to their sales with a one-period lag. Of the additional $10 billion of income, one fourth ($2.5 billion) is used to pay additional taxes, and one fifth of the remainder ($1.5 billion) goes into saving while the other four fifths ($6 billion) is spent on consumption. Investment spending remains at the new level of $130 billion in the fourth period, and with the added $6 billion of consumption expenditures, total sales rise to $816 billion. Accordingly, output and income rise to $816 billion in the fifth period, two periods after the rise in investment began. The student should trace out for himself the further increases in income and output that occur in periods six and seven. (The material in the last three lines of Table 7–1 should be disregarded for the present. It will be referred to later.)

The static multiplier for an autonomous change in investment, which we discussed earlier, is given by the expression

$$\frac{\Delta Y}{\Delta I} = \frac{1}{1 - c(1 - x)} = \frac{1}{1 - 0.8(1 - 0.25)} = 2.5$$

Thus the new equilibrium income corresponding to the increased rate of investment is above the old equilibrium level by $10 billion times the multiplier of 2.5, or by $25 billion. That is, the new equilibrium is $825 billion, as shown in the last column of Table 7–1.

Stability of Equilibrium

The conditions under which the dynamic process we have been describing will actually settle down to the new equilibrium determined by the static multiplier will now be examined. In the example worked out above (see Table 7–1), the marginal propensity to consume *out of GNP* is three fifths. That is, of an addition to GNP, one fourth is taken in taxes, while four fifths of the remaining three fourths is spent on consumption. Thus, if c_y is the marginal propensity to consume out of GNP, we have

$$c_y = c(1 - x)$$

or, in the specific example,

$$c_y = \tfrac{4}{5}(1 - \tfrac{1}{4}) = (\tfrac{4}{5})(\tfrac{3}{4}) = \tfrac{3}{5}$$

TABLE 7-1

Illustration of Dynamics of the Multiplier
(billions of dollars)

Variable	Definition	How Calculated	Time Period: Original Equilibrium 1	2	3	4	5	6	7	New Equilibrium
Y_t	GNP	$C_{t-1} + I_t + G_{t-1}$	800	800	800	810.0	816.0	819.60	821.760	825.00
T_t	Taxes	$0.25\ Y_t - 40$	160	160	160	162.5	164.0	164.90	165.440	166.25
Y_t^d	Disposable income	$Y_t - T_t$	640	640	640	647.5	652.0	654.70	656.320	658.75
C_t	Consumption	$0.8Y_t^d$	512	512	512	518.0	521.6	523.76	525.056	527.00
I	Planned investment	Autonomous	120	120	130	130.0	130.0	130.00	130.000	130.00
G	Government purchases	Autonomous	168	168	168	168.0	168.0	168.00	168.000	168.00
S_t	Saving	$Y_t^d - C_t$	128	128	128	129.5	130.4	130.94	131.264	131.75
I_t^n	Inventory investment	$Y_t - (C_t + I_t + G_t)$	0	0	−10	−6.0	−3.6	−2.16	−1.296	0
I_t^a	Actual investment	$I + I_t^n$	120	120	120	124.0	126.4	127.84	128.704	130.00
B_t	Government surplus or deficit (−)	$T_t - G_t$	−8	−8	−8	−5.5	−4.0	−3.10	−2.560	−1.75

Referring to Table 7–1, the amount by which sales in period five (output in period six) exceeds sales in period two (output in period three) can be computed as follows:

Increased investment spending in period 5	10.0
Consumption in period 5 resulting from respending of income generated by 10 of increased investment spending in period 4 (.60 × 10)	6.0
Consumption spending in period 5 resulting from second round of respending of income generated by 10 of increased investment spending in period 3 (.60 × .60 × 10)	3.6
Total	19.6

Thus, by the expiration of three periods after the multiplier process begins, income has risen above its original level by 19.6

In general, by the same reasoning used in this example, if ΔY_n is the amount by which income in the *n*th period after the increase in investment (ΔI) began exceeds the original equilibrium level, we have

$$\Delta Y_n = \Delta I + c_y \Delta I + c_y^2 \Delta I + \cdots + c_y^{n-1} \Delta I$$

This is the summation of a geometric progression having n terms, with the first term ΔI and common ratio c_y. According to the formula for the sum of n terms of a geometric progression which appears in any elementary algebra book, we have[5]

$$\Delta Y_n = \frac{1 - c_y^n}{1 - c_y} \Delta I$$

If c_y (the marginal propensity to consume out of GNP) is less than unity, c_y^n will dwindle away toward zero as n (the time period considered) grows longer. Thus, as n increases, the value of ΔY, the rise in income since the beginning of the expansion, approaches

$$\Delta Y = \frac{1}{1 - c_y} \Delta I$$

[5] Actually the derivation is quite simple. We have

$$\Delta Y_n = \Delta I + c_y\Delta I + c_y^2\Delta I + \ldots + c_y^{n-1}\Delta I \qquad \text{(a)}$$

Multiplying both sides of this equation by $-c_y$, we obtain

$$-c_y\Delta Y_n = -c_y\Delta I - c_y^2\Delta I - \ldots - c_y^{n-1}\Delta I - c_y^n\Delta I \qquad \text{(b)}$$

Adding (a) and (b) yields

$$\Delta Y_n - c_y\Delta Y_n = \Delta I - c_y^n\Delta I$$

This easily simplifies to

$$\Delta Y_n = \frac{1 - c_y^n}{1 - c_y}\Delta I$$

or, since $c_y = c(1 - x)$,

$$\Delta Y = \frac{1}{1 - c(1 - x)} \Delta I$$

This is, of course, the static multiplier shown earlier.

If the marginal propensity to consume out of GNP is greater than unity, income keeps on expanding indefinitely at a geometric rate and does not settle down to the level determined by the static multiplier. Thus, when the marginal propensity to consume out of GNP is greater than unity, the system is said to be *unstable.* A commonsense way of looking at this is that if the marginal propensity to consume is greater than unity, when income rises, consumption spending rises more than income—that is, consumers borrow or dip into past savings to finance a part of current spending. Thus, saving falls as income rises. Since an increase in investment calls for an increase in saving to match it, in this case an increase in investment *lowers* the equilibrium level of income. However, the added investment spending stimulates production and actually raises income. Thus, income *moves away* from the new equilibrium rather than toward it; moreover, saving falls as income rises so that the gap between investment and saving gets progressively larger, thus continuing to drive income up at ever-increasing speed.[6]

The Speed of the Multiplier

Assuming that c_y, the marginal propensity to consume out of GNP, is less than unity so that the stability condition is satisfied, income will approach the new equilibrium gradually, or *asymptotically;* that is, it will never quite get to the equilibrium level but will approach more and more closely to it period by period. In the example presented in Table 7–1, income approaches the new equilibrum fairly rapidly; it is getting close—in fact, it has gone almost 80 percent of the way—by the third period after the initial rise in investment (period six). Since the static multiplier takes an indefinitely long time to unfold, it is really of only academic interest. More significant in most cases is the extent of the multiplier effect which will be realized in, say, a year's time. A much longer period than that is not likely to be very meaningful, since further autonomous changes in investment or government expenditures are likely to supervene before the more distant effects of the initial change are realized. The multiplier applicable to a specific finite number of time periods is sometimes called the *truncated multiplier.* Thus, in the above example, the truncated multiplier for three

[6] If c_y is greater than unity, the static multiplier is negative; for example, if $c_y = 1.1$, the multiplier is -10. This is often puzzling to students, but it merely reflects the fact that in such a case it is necessary to lower income in order to raise saving, so that a rise in investment reduces *equilibrium* income. It is suggested that the student work out a case of this kind graphically, using a 45° guideline diagram such as that employed in Figure 6–2.

periods is 1.96, since a rise in investment of $10 billion per period beginning in period three raises income by $19.6 billion three periods later (in period six). The speed with which the multiplier works, measured in calendar time, depends on the length of the multiplier time period and on the size of the marginal propensity to consume out of GNP.

The length of the *multiplier time period*—i.e., the unit time period in the multiplier process—is primarily a question of the time required to adjust production to sales.[7] For some kinds of services, such as medical care, this period is very short, since the services are produced in direct and immediate response to demand. On the other hand, in the case of many types of consumer goods, the production process involves several stages, and the goods are initially supplied by drawing down inventory stocks. For example, if there is an increase in the demand for shoes, the demand will be met in the first instance by drawing down the inventories of retail dealers. As inventories become depleted, retailers will place increased orders with wholesalers or perhaps directly with manufacturers. Wholesalers in turn will expand their orders with manufacturers as their inventories are reduced. Manufacturers will increase their rates of production as their stocks of goods are drawn down and will also place increased orders for raw materials, thus stimulating production at the raw material stage.

Some additional income, chiefly in the form of business profits, will be generated at the retail and wholesale level, but the major increase will be at the manufacturing stage, where the increase in production and employment will generate additional wage income. Thus, most of the income generation will occur only after a time lag which involves the reduction of inventories at the retail and wholesale levels, the placing of orders at the various stages, the increase in production rates at the manufacturing stage, and the completion of the process of production.

The length of the multiplier time period will therefore depend upon such factors as the time it takes for retailers and wholesalers to recognize the increase in sales as a semipermanent phenomenon and not merely a temporary random aberration and to react to it by placing additional orders, the time it takes manufacturers similarly to recognize a significant increase in sales and to adjust their production rates accordingly, and so on.

It is extremely difficult to estimate the length of the time period with any great degree of confidence, but the general order of magnitude is probably something like three to four months. No doubt the period varies with business conditions; for example, in a time of severe depression, inventories are likely to have piled up as a result of a decline in sales so that when sales

[7] For a more extended discussion of this subject, see Gardner Ackley, "The Multiplier Time Period: Money, Inventories, and Flexibility," *American Economic Review*, Vol. XLI, June 1951, pp. 350–68, reprinted in W. L. Smith and R. L. Teigen (eds.), *Readings in Money, National Income, and Stabilization Policy* (rev. ed.; Homewood, Ill.: Richard D. Irwin, Inc., 1970).

do pick up, firms at the retail, wholesale, and manufacturing stages are likely to be content to take the opportunity to work off redundant inventories and to postpone for a considerable time the placing of additional orders and the acceleration of production.

The other factor that determines the speed of the multiplier is the *marginal propensity to consume out of GNP.* Table 7–2 shows values of the

TABLE 7–2

Comparison of Static Multipliers and Truncated Multipliers for Four Periods with Alternative Values of the Marginal Propensity to Consume out of GNP

(1) *Marginal Propensity to Consume out of GNP* (c_y)	(2) *Static Multiplier*[a]	(3) *Truncated Multiplier for 4 Periods*[b]	*Percent of Static Multiplier Realized in 4 Periods* [(3) ÷ (2)]
0.9	10.00	3.439	34.4
0.8	5.00	2.952	59.0
0.7	3.33	2.533	76.0
0.6	2.50	2.176	87.0
0.5	2.00	1.875	93.8
0.4	1.67	1.624	97.4
0.3	1.43	1.417	99.1
0.2	1.25	1.248	99.8

[a] Static multiplier $= \frac{1}{1 - c_y}$.

[b] Truncated multiplier for four periods $= 1 + c_y + c_y^2 + c_y^3$.

static multiplier and of the four-period truncated multiplier corresponding to values of c_y, the marginal propensity to consume out of GNP, between 0.2 and 0.9. As can be seen from the table, both the four-period multiplier and the static multiplier are larger the greater is the marginal propensity to consume. However, the multiplier unfolds more rapidly for low values of the marginal propensity to consume than for high values. For a marginal propensity to consume of 90 percent, only 34.4 percent of the ultimate static multiplier effect has made its appearance by the fourth period. This percentage rises rapidly as the marginal propensity to consume declines; with marginal propensities to consume of 50 percent or lower, more than 90 percent of the ultimate static multiplier effect has been achieved by the fourth period.

Let us consider the speed of the multiplier for the relatively complex model developed in the concluding section of Chapter 6, which allows for changes in corporate profits and dividends and also takes account of a variety of taxes and transfer payments. For the numerical illustration we developed (see pages 126–33 of Chapter 6) the marginal propensity to consume out of GNP was 0.57825, and the corresponding value of the static multiplier was 2.37. If we assume that the multiplier time period is one quarter, the four-quarter multiplier will be 2.11.[8] In this case, the four-period multiplier is 89 percent of the static multiplier. That is, by the end of four quarters, approximately nine tenths of the total multiplier effect has been achieved. By similar reasoning, the four-quarter multiplier for a change in personal taxes would be 89 percent of the corresponding static multiplier. In our numerical example, the static multiplier for a change in personal taxes was −2.13 (see page 000 of Chapter 6); accordingly the one-year multiplier for a personal tax change would be −1.90 (−2.13 × 0.89 = −1.89057).

EQUALITY OF SAVING AND INVESTMENT

In the introduction to this part of the book, the distinction between planned (ex ante) and actual (ex post) saving and investment was made and its implications for economic analysis were discussed briefly. At this point it will be useful to return to this matter and examine it a little more carefully.

The easiest way to clarify the distinction between planned and actual magnitudes is to take a specific example. Suppose the Atlas Furniture Company, finding that it needs additional plant capacity to meet the demands of customers, buys $100,000 of new machinery from the Brown and Wilson Machinery Corporation. For simplicity, suppose that the machinery ordered is of a standard type, and Brown and Wilson fills the order by shipping machinery from stock. Suppose further that the machinery shipped is valued in Brown and Wilson's inventory at $84,000, the amount it cost the company to produce it. The following table shows the contributions of the two firms to national income, consumption, saving, and investment resulting from the transaction.

Contribution to:	*Atlas Furniture Co.*	*Brown & Wilson Machinery Corp.*	*Total*
Y...........	.....	+16,000	+16,000
C...........	.....	.....	.....
I...........	+100,000	−84,000	+16,000
S...........	.....	+16,000	+16,000

The purchase of $100,000 of machinery by the Atlas Furniture Company is clearly investment. As a result of selling machinery out of its inventory, the Brown and Wilson Machinery Corporation has inventory disinvestment of $84,000. By selling goods costing $84,000, at a price of $100,000, Brown and Wilson realizes a profit of $16,000, which is an addition to national

[8] The calculation is as follows: four-period multiplier $= 1 + c_y + c_y^2 + c_y^3$. If c_y is 0.57825, this becomes $1.00000 + 0.57825 + 0.33437 + 0.19335 = 2.10597$.

income. Since this income is not immediately distributed to stockholders in dividends, its immediate effect is to add $16,000 to gross saving in the form of corporate retained earnings. Thus, gross investment as measured in the national income accounts has increased by $16,000, the difference between Atlas's investment in plant and equipment and Brown and Wilson's disinvestment in inventories. Gross saving has also increased by $16,000, the increase in Brown and Wilson's retained earnings. The transaction therefore affects investment and saving equally, increasing each by $16,000.[9]

However, the different elements of saving and investment have different significance for economic analysis. The $100,000 investment in plant and equipment by Atlas may be called *planned* investment in the sense that Atlas obviously decided that it needed the additional equipment and deliberately set out to acquire it. On the other hand, the $84,000 disinvestment in inventories by Brown and Wilson is, in a sense at least, *unplanned* disinvestment, since Brown and Wilson merely responded to Atlas's order by shipping the goods out of stock. Under normal circumstances, assuming that the company held the size and composition of inventories that it desired before the sale, it can be expected to produce the equipment necessary to replenish its inventory, and this production will generate an equivalent amount of income. There is, of course, no guarantee of this; if Brown and Wilson's sales have been falling off, with goods unsold and inventories piling up, the company may welcome this opportunity to reduce redundant inventory stocks. Thus, as a result of a decision not to replenish depleted stocks, all or part of the reduction in inventories, which from the point of view of the transaction we are considering is more or less automatic *unplanned* disinvestment, may become *planned* disinvestment.

In any case, it may be noted, it is the planned portion of the investment that stimulates production and income. If Brown and Wilson fully replaces the inventory, Atlas's purchase of equipment will exert its full effect on production and income; to the extent that Brown and Wilson permits its inventory to decline without replenishment (so that all or part of the initial unplanned disinvestment becomes planned disinvestment), the stimulus to production and income will be reduced. Consequently, it is changes in planned investment that tend to drive the level of production, income, and employment up or down in accordance with the analysis presented in this chapter.

A similar distinction can be made in connection with saving. The $16,000 saving by Brown and Wilson may be regarded, in relation to the immediate transaction giving rise to it, as *unplanned* saving in the sense that the company's profits have increased by $16,000, and immediately after the transaction there has not yet been time for the company to pay any of it out in dividends and thereby permit stockholders to spend a portion on con-

[9] The reader will note that saving and investments are not equal for each of the participants in the transaction taken separately, but only for the two taken together—that is, for the economy as a whole.

sumption. In the normal course of events, it is likely that a part of the $16,000 will be paid over in dividends to stockholders and that some of the dividend income will be spent on consumption. To the extent that this happens, it will tend to stimulate further production, income, and employment, constituting part of the multiplier process which we have been discussing.

To the extent that Brown and Wilson *decides* to keep back a portion of the $16,000 as retained earnings and stockholders *decide* to save a portion of the dividend income paid to them out of the $16,000, there will be planned saving. It is this portion of saving that acts as a check to income expansion by withholding dollars from the income stream. The portion of saving that is unplanned will be dissaved—that is, spent on consumption—within a relatively short period of time and will contribute to the continued expansion of income.

Let us take a couple of other illustrative transactions. Suppose that Henry Wilkinson buys a television set from General Appliances, Inc., for $175. Assume that the set cost General Appliances $125, so that the markup (or profit margin) on the transaction is $50. The results are shown in the following table.

Contribution to:	*Henry Wilkinson*	*General Appliances, Inc.*	*Total*
Y	...	+50	+50
C	+175	...	+175
I	...	−125	−125
S	−175	+50	−125

It is apparent that, once again, the transaction affects saving and investment equally for the economy as a whole, reducing each by $125. We leave it to the reader to consider the problem of disentangling the planned and unplanned portions of saving and investment in this case.

If we introduce government, the analysis becomes a little more complex, but the principles are not changed. Suppose, for example, that James Larson pays his federal income tax in the amount of $1,200. The results are as follows:

Contribution to:	*James Larson*	*Federal government*	*Total*
Y	....	....	....
C	....	....	....
I	....	....	....
S	−1,200	....	−1,200
Tax receipts	....	+1,200	+1,200
Government expenditures	....	....	....
Government surplus	....	+1,200	+1,200

It is apparent that in the national accounts, the relationship

Gross investment = Gross saving + Government surplus

which replaces the simple identity of saving and investment when government is introduced, is satisfied, since saving is reduced by $1,200 while the surplus is increased by the same amount. The reader may consider the effect on production and income that is likely to result from the fact that part of the $1,200 of dissaving, which is a necessary immediate statistical by-product of the tax payment, may, in effect, be unplanned.

As a final example, suppose the city of Stamford, Connecticut, pays out $15,000 in wages to workers employed in the making of street repairs. The results would be as follows:

Contribution to:	*Highway Workers*	*Stamford City Government*	*Total*
Y	+15,000	.....	+15,000
C	.....	.....	.....
I	.....	.....	.....
S	+15,000	.....	+15,000
Tax receipts	.....	.....	.....
Government expenditures	.....	+15,000	+15,000
Government surplus	.....	−15,000	−15,000

Once again, the identity,

Gross investment = Gross saving + Government surplus

is satisfied in the national income accounts, since saving is increased by $15,000, while the surplus is reduced by the same amount. However, there is likely to be a stimulus to the economy resulting from the fact that a portion of the $15,000 of saving is almost certainly unplanned, resulting merely from the fact that immediately after the payment the workers have not had time to spend the desired portion of their incomes.

It should be apparent by now that the effects of *any single* transaction are necessarily such as to satisfy the condition,

Actual gross investment = Actual saving + Government surplus

If this condition is true for every transaction separately, it must be true for any combination of transactions. Consequently, the condition must be satisfied for the economy as a whole during *any time period whatever,* whether it be five minutes or half a century.

We can summarize the results of the above analysis as follows:

1. Conceptually, it is possible to distinguish between (*a*) planned investment, (*b*) unplanned inventory investment, and (*c*) actual investment as measured in the national income accounts. Similarly, a distinction can be

made in principle between (*a*) planned saving, (*b*) unplanned saving, and (*c*) actual saving as shown in the national income accounts. Thus,

Planned investment + Unplanned investment = Actual investment

and

Planned saving + Unplanned saving = Actual saving

2. Equilibrium of the level of income and employment in a closed economy requires that the following condition be satisfied:

Planned gross investment = Planned saving + Government surplus

When planned investment exceeds planned saving plus government surplus, the level of income and employment (and/or prices) will rise; when planned saving plus government surplus exceeds planned investment, income and employment (and/or prices) will fall.

3. If during any period planned investment is not equal to planned saving plus government surplus, there must necessarily be just enough unplanned inventory investment and/or unplanned saving so that

Actual investment = Actual saving + Government surplus

Consequently, this condition, which is reflected in the national income accounts, is always satisfied, no matter what the state of the economy is.

This analysis is related to the question of lags in the circular flow of income discussed earlier in this chapter. Referring to the model involving an output lag but no expenditure or earnings lags, as illustrated in Table 7–1, it is clear that there can be no unplanned saving, since consumption (and therefore saving) is adjusted to income within each period. Consequently, planned saving and actual saving are always equal; i.e., saving plans are always fulfilled. However, there can be differences between planned and actual investment, with unplanned inventory investment making up the differences. In the example of Table 7–1, for instance, in period four, total output (Y_t) is \$810 billion, while total sales ($C_t + I + G$) amount to \$816 billion. Clearly, the only way in which producers could sell \$816 billion of goods and services while producing only \$810 billion is to draw down their inventories by \$6 billion. Thus, there is inventory disinvestment (designated by I_t^u) of \$6 billion, as shown in the "Inventory investment" entry in Table 7–1. When this unplanned disinvestment is subtracted from the \$130 billion of planned investment (designated as I), the result is \$124 billion of actual investment (designated as I_t^a). This amount is equal to saving of \$129.5 billion minus the government deficit of \$5.5 billion. These relations hold true in each period. When the economy is in equilibrium (in the first two periods and the last period shown in Table 7–1), unplanned inventory investment (I_t^u) is, of course, zero, and planned investment (I) is equal to actual investment (I_t^a).

Introduction of an expenditure lag and/or earnings lag permits the occurrence of unplanned saving. The reason for this is that with such lags, planned saving is determined by the income of an earlier period, while actual saving is calculated by subtracting consumption from the current period's income. Thus, if current income differs from past income, actual saving will depart from planned saving so that there must necessarily be unplanned saving.

INVENTORY ADJUSTMENTS

Changes in inventory investment are so closely related to the multiplier that they may actually be considered a part of the multiplier process. This can be seen from the simple illustration of the dynamic multiplier employing an output lag shown in Table 7–1. In a sense, the multiplier process "feeds on" inventories in this model. Beginning with period three—the period in which the autonomous increase in investment takes place—inventories decline steadily, that is, inventory investment (I_t^n) is consistently negative. In fact, production and income increase each period by as much as inventories declined the period before, so that the cumulative decline in inventory stocks is exactly equal to the rise in income. For example, three periods after the autonomous increase in investment, income has risen to $819.6 billion or $19.6 billion above its original level, and the combined fall in inventories during these three periods (periods three, four, and five) is precisely $19.6 billion. When the new equilibrium is reached, income will be $25 billion above its original level, inventory stocks will have been reduced by $25 billion, and current investment in inventories (I_t^n) will be zero.

This pattern of behavior is obviously quite unrealistic. Businessmen can hardly be expected to look with equanimity upon a progressive decline in their inventory stocks under normal circumstances. They will surely have some kind of inventory policy which will cause them to take measures to maintain their inventories at a suitable level or perhaps increase them as sales rise.

A Model with Inventory Adjustment

When allowance is made for inventory adjustments on the part of businessmen, the dynamic multiplier process of income change is altered rather significantly. The effects are illustrated in Table 7–3, which shows the working of the multiplier when the marginal propensity to consume is 50 percent and when businessmen are assumed to follow the very simple policy of attempting to maintain constant inventory stocks. For simplicity we assume in this case that there are no government expenditures or taxes.

TABLE 7-3

Dynamic Multiplier Sequence with Inventory Adjustment
Based on Constant Desired Inventory Stock

Variable	Meaning	How Calculated	Time Period: Original Equilibrium 1	2	3	4	5	6	7	8	New Equilibrium
Y_t	Gross national product	$C_{t-1} + I^f_{t-1} + D_{t-1}$	440.0	440.0	440.0	460.0	470.0	470.0	465.0	460.0	460.0
C_t	Personal consumption expenditures	$0.5Y_t + 100$	320.0	320.0	320.0	330.0	335.0	335.0	332.5	330.0	330.0
I_t^f	Fixed investment	Autonomous	120.0	120.0	130.0	130.0	130.0	130.0	130.0	130.0	130.0
S_t	Gross saving	$Y_t - C_t$	120.0	120.0	120.0	130.0	135.0	135.0	132.5	130.0	130.0
I_t^n	Actual inventory investment	$Y_t - (C_t + I_t^f)$	0.0	0.0	−10.0	0.0	5.0	5.0	2.5	0.0	0.0
I_t^a	Actual total investment	$I_t^f + I_t^n$	120.0	120.0	120.0	130.0	135.0	135.0	132.5	130.0	130.0
K_t^n	Actual inventory stocks end of period	$K^n_{t-1} + I_t^n$	100.0	100.0	90.0	90.0	95.0	100.0	102.5	102.5	100.0
K_t^d	Desired inventory stocks	Autonomous	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
D_t	Demand for inventories	$K_t^d - K_t^n$	0.0	0.0	10.0	10.0	5.0	0.0	−2.5	−2.5	0.0

The consumption function employed in this model is

$$C_t = 0.5Y_t + 100$$

and the system is assumed to be in equilibrium initially, with autonomous investment in fixed capital of $120 billion per period. It is further assumed that businessmen desire to hold inventory stocks of $100 billion; initially actual stocks are at this level, so that the economy is in equilibrium with income at a level of $440 billion (as shown in the first two columns of Table 7–3).

In period three autonomous fixed investment (I') rises to $130 billion and remains at that level. As a result of the increase in investment, total sales rise to $450 billion in period three, and since production in that period is only $440 billion, inventories are depleted by $10 billion. As a result, an additional demand for goods of $10 billion arises out of the desire of businessmen to replace depleted inventories. Consequently, production and income rise to $460 billion in period four; in addition to producing an amount equal to the $450 billion of sales in period three, firms produce a further $10 billion to make good the previous period's depletion of inventories. However, the $20 billion rise in income in period four causes consumption to rise by $10 billion; this uses up the $10 billion produced for the purpose of restoring inventories to their normal level. As a result, inventories remain at $90 billion, $10 billion short of the desired level. Thus, the $10 billion demand for inventories is recreated in period four as businessmen once again attempt to rebuild their stocks, and income rises to $470 billion in period five. However, this rise in income causes consumption to rise by $5 billion, thus using up half of the $10 billion produced for inventories, so that inventories rise only to $95 billion. With inventories still $5 billion short of the target, a further inventory demand of $5 billion is generated.

Since consumer demand in period five is $5 billion more than in the previous period, while inventory demand is $5 billion less, total demand remains at $470 billion and income is again $470 billion in period six. At this level, aggregate demand for consumption and investment is only $465 billion, so that inventories rise by $5 billion to the desired level of $100 billion. However, despite the coincidence of desired and actual inventories, income cannot remain at this level, since inventory demand now drops to zero while consumption and investment demand are only $465 billion, which is not sufficient to maintain income at $470 billion. Consequently, income drops to $465 billion in period seven. As a result of the fall in income, demand drops to $462.5 billion and inventories accumulate to the extent of $2.5 billion, causing stocks to rise above the desired level of $100 billion. Thus, inventory demand becomes negative, causing income to drop further in period eight. As the student can verify for himself by carrying the computations of Table 7–3 through a few more periods, the decline in income will eventually reverse itself. It should be apparent by now that the initial rise in fixed investment has started a self-sustaining cyclical fluctuation in income.

The last column of Table 7–3 shows the equilibrium position corresponding with investment of $130 billion per period. In this model, the *static* multiplier is equal to the reciprocal of the marginal propensity to save, just as in the case of our earlier model without inventory adjustment. Since the marginal propensity to save is 0.5, the multiplier is 2, and the initial $10 billion increase in investment causes equilibrium income to rise by $20 billion to $460 billion. However, it should be noted that in this model two conditions must simultaneously be satisfied in order for equilibrium to be achieved: (1) income must be $460 billion, and (2) inventories must be at the desired level of $100 billion. In period four and in period eight, income is $460 billion, but in neither case can it remain there, because actual inventory stocks do not coincide with the desired level.

Inventories and the Multiplier

Although the model used in the construction of Table 7–3, with its assumption that the desired level of inventories is constant, is unquestionably a great oversimplification, it does nevertheless illustrate two features that actually seem to characterize the mechanism of short-run income adjustment in the real world.

1. The existence of an inventory reaction, even when based upon the simple assumption of a constant desired inventory stock, introduces a self-sustaining cyclical movement into the economy. Thus, when the conditions of equilibrium are disturbed by a rise in autonomous expenditures, income first rises above the new equilibrium level, then begins to decline, falls below equilibrium, rises again, and so on. In the illustration of Table 7–3, the fluctuations are *damped*—that is, they get progressively smaller and income approaches more and more closely to the new equilibrium level as it fluctuates around that level.[10] The time path of income is of the sort shown in Figure 7–1B. This contrasts with the smooth approach to the new equilibrium which characterizes the simple multiplier without inventory adjustment, as typified by the path shown in Figure 7–1A.

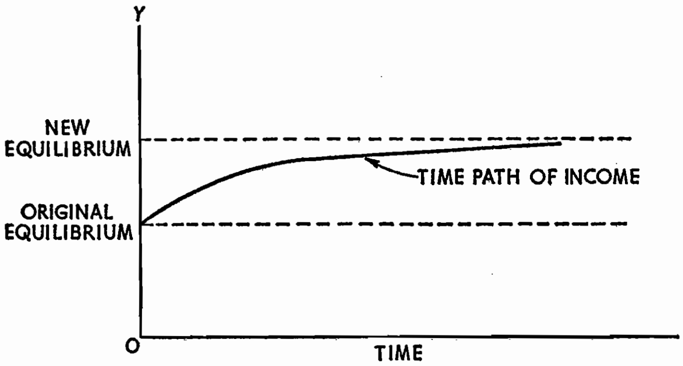
2. The presence of an inventory adjustment makes the multiplier larger for the first few periods after the initiating change in autonomous expenditures than it would be in the absence of such an adjustment. In the example shown in Table 7–3, the increase in investment of $10 billion causes income to rise $30 billion above its original level at the expiration of three periods (in period six). Thus, the multiplier including inventory adjustment for

---

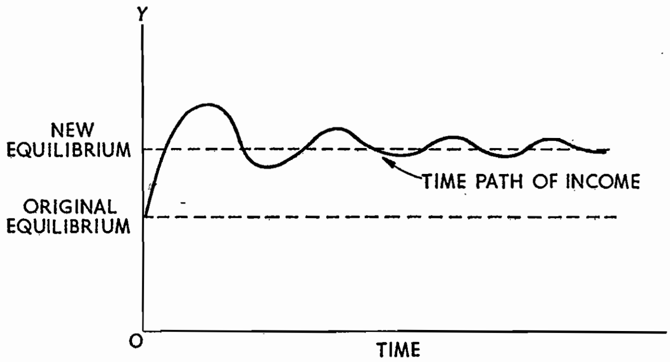
[10] This particular model will generate damped fluctuations provided the marginal propensity to consume is less than unity.

FIGURE 7–1

A. Multiplier Process without Inventory Adjustment



B. Multiplier Process with Inventory Adjustment Based on Constant Desired Inventories



three periods is 3 in this case. The student can easily verify that with a marginal propensity to consume of 50 percent, in the absence of inventory adjustment, an increase in investment of $10 billion would cause income to rise by $17.5 billion at the end of three periods; that is, the three-period multiplier would be 1.75. Thus, the presence of even such a simple type of inventory reaction increases the three-period multiplier by about 70 percent.

There can be little doubt that inventory adjustments in practice are much more complex than those involved in the simple model used in Table 7–3. Inventories of finished goods are held by merchants, wholesalers, and manufacturers in order to have sufficient quantities and varieties of goods on hand to meet customers' orders without delay. Stocks of raw materials are needed to provide assurance against interruptions in production due to possible delays in the filling of orders by suppliers and to permit economies associated with periodic ordering and bulk purchasing. Inventories of goods in process of production are to a considerable extent determined by the technology of the production process, together with the current rate of production.[11] As a result, the size of desired inventory stocks held for all of these purposes tends to rise and fall with the rate of sales or production. Thus, as sales rise, firms not only speed up their rates of production in an attempt to prevent their inventories from falling, as was assumed in constructing Table 7–3, but in addition they may speed up production still more in order to increase their inventories more or less in pace with sales. This kind of behavior accentuates the tendency for inventory adjustments (*a*) to enlarge the size of the multiplier applicable to the earlier periods of income change and (*b*) to introduce self-perpetuating cyclical movements into the economy.[12]

While there can be little doubt that inventory adjustments tend to amplify short-run movements in income set off by changes in fixed investment and government expenditures, the state of our knowledge concerning such adjustments is not entirely satisfactory. However, some estimates are available relating to the multiplier including inventory adjustments. Calculations based upon the econometric model of the United States economy developed in the Research Seminar in Quantitative Economics at the University of Michigan indicate that the one-year multiplier applicable to changes in business expenditures for plant and equipment or in purchases of goods and services inclusive of inventory reactions is in the neighborhood of 2.[13] Since this model uses annual data, the result means that an increase in government purchases at the beginning of a year will raise the GNP for that

---

[11] As a simple illustration of this point, suppose a firm is producing 10 units of a certain product per day, and suppose further that 30 days are required to complete the process of production. If production is proceeding steadily, the firm will start production on 10 units each day and at the same time 10 units will be completed each day. At all times there will be in the production process 10 units started on each of the last 30 days, or a total of 300 units. On the average these units will be about half finished, so that the inventory of goods in process of production will consist of the equivalent of about 150 units of the product. Assuming 240 working days per year, annual output will be 2,400 units, and inventory stocks will be approximately 1/16 of annual output.

[12] For an extensive discussion of the cyclical effects of inventory adjustments, see L. A. Metzler, "Nature and Stability of Inventory Cycles," *Review of Economic Statistics,* Vol. XXIII, 1941, pp. 113–29; "Factors Governing the Length of Inventory Cycles," *Review of Economic Statistics,* Vol. XXIX, February 1947, pp. 1–15; and "Business Cycles and the Modern Theory of Employment," *American Economic Review,* Vol. XXXVI, June 1946, p. 278.

[13] For a description of the Michigan model, see D. B. Suits, "The Economic Outlook in 1968," in *The Economic Outlook for 1968,* Papers presented at the Fifteenth Annual Conference on the Economic Outlook at the University of Michigan (Ann Arbor: Department of Economics, University of Michigan, 1968), pp. 1–31. An earlier version of the model including the multipliers applicable to various exogenous changes is described in D. B. Suits, "Forecasting and Analysis with an Econometric Model," *American Economic Review,* Vol. LII, March 1962, pp. 104–32, reprinted in R. A. Gordon and L. R. Klein (eds.), *Readings in Business Cycles* (Homewood, Ill.: Richard D. Irwin, Inc., 1965), pp. 597–625.

year by an amount equal to about twice the original increase in purchases. Another model, which makes use of quarterly rather than annual data and deals with the functioning of the economy during recessions, indicates a four-quarter multiplier of about 2.5, including inventory reactions.[14]

[14] J. S. Duesenberry, Otto Eckstein, and Gary Fromm, "A Simulation of the United States Economy in Recession," *Econometrica*, Vol. XXVIII, October 1960, Table V, p. 764. (This table is not included in the reprinted version of the paper in Klein and Gordon (eds.), *Readings in Business Cycles, op. cit.*) It may be noted that the multiplier for the *year*, following a change in, say, private investment, will generally be different from the multiplier applicable to the *fourth quarter* after the change. For example, taking a period to be a quarter in Table 7–1, the increase of $10 billion in investment in period 3 will raise income for the ensuing year by $16.84 billion (the average value of GNP for periods 4, 5, 6, and 7 is $816.84 billion). Thus, the multiplier for the year is 1.684. However, the multiplier applicable to the fourth quarter after the increase in investment began (period 7) is 2.176 since GNP for that period is $821.76 billion.

Chapter 8

INVESTMENT I: THE MARGINAL EFFICIENCY OF INVESTMENT

In our discussion of the multiplier in the previous two chapters, we treated investment expenditures (other than for inventories) as an *autonomous variable*—that is, we took such expenditures as given in particular situations and considered what would happen to the level of income when they changed, but we made no effort to consider the factors determining them. In this and the next two chapters we shall extend the scope of our model of the income determination mechanism to cover some of the determinants of investment expenditures.

SCOPE OF ANALYSIS

We shall use the word investment in its basic economic meaning of *real investment*—that is, the *accumulation of real capital goods.* We shall exclude investment in inventories from our discussion, since the behavior of inventory investment was discussed in the previous chapter in connection with the multiplier. Moreover, we shall consider only private investment by profit-oriented economic units, omitting investment by governmental units in such things as schools, dams, and highways. Thus, investment in the present context means expenditures by private business firms on durable equipment and on the construction of plants, stores, and other commercial buildings.

Gross and Net Investment

In principle, it is possible to distinguish between new investment—additions to the stock of real capital—and replacement investment—investment activity designed to maintain the existing stock in the face of wear and tear, obsolescence, and destruction by fire, flood, or other catastrophe. *Net investment* includes only new investment in the above sense, while *gross investment* includes both new investment and replacement investment.

As indicated below, gross investment is the more appropriate concept for some purposes, while in other situations net investment is more suitable.

However, in practice it is rather difficult to measure net investment. The most obvious way to obtain an estimate of net investment is to deduct depreciation allowances from gross investment; however, since depreciation allowances are usually based upon the original cost of individual plants and pieces of equipment which were purchased in different years when prices were at various levels, it is difficult to obtain a meaningful estimate of net investment in this way. In addition, there is a more fundamental difficulty which arises from the fact that a worn-out piece of equipment is rarely replaced by another exactly like it; rather it is usual to replace discarded equipment with the latest and most efficient type available. For this reason, in most cases replacement of worn-out or obsolescent equipment increases productive capacity to some extent and therefore includes an element of new investment. Hence it becomes *conceptually* difficult to disentangle replacement investment from new investment.

Investment as a Component of Aggregate Demand

In this chapter, we shall be concerned with private investment expenditures as a part of aggregate demand for final output. In this context, gross investment is a more appropriate concept than net investment. This is true because one of our more important objectives is to uncover the factors determining levels of employment and prices. Employment is generated by the production of capital goods whether they are intended for new investment or for replacement; and when there is pressure on the existing capacity of the economy, an increase in the demand for capital goods whether for new investment or for replacement may exert upward pressure on the price level. Since allowances for depreciation constitute a deduction from the gross income flow and in any particular period may not—in fact, commonly will not—exactly equal expenditures for replacement of depreciated equipment, an excess of depreciation allowances over replacement expenditures may be a contractive influence on economic activity. Thus, it is best to consider the relation between *gross investment* (including replacement) and *gross saving* (including depreciation) for our present purposes.

Investment as a Source of Productive Capacity

In addition to being an important component of final demand for current output, investment also adds to the productive capacity of the economy, both by increasing the size of the capital stock and by serving as the vehicle through which new technology is introduced. There are many important and somewhat controversial issues concerning the role of capital formation in economic growth, which will be taken up at length later in this book.[1] In

[1] See Part IV.

this chapter, however, these issues are almost entirely neglected, and the emphasis is placed on investment as a component of aggregate demand. That is, gross investment is taken into account as an element of final demand, but the capacity-creating effects of investment are not allowed for. This means that the analysis has limitations which may make it unsuitable in some situations, a fact that should be borne in mind by those who use it. However, for purposes of short-term analysis of aggregate demand, which is our present concern, failure to take account of the effects of investment on productive capacity is not likely to be a source of serious distortion.[2]

THE MARGINAL EFFICIENCY CONCEPT

Presumably business decisions to invest in additional productive facilities —to install new machinery, build new factories, stores, office buildings, and so on—are based on some kind of profit maximization calculus. That is, investment takes place only when those responsible for decision making believe that it will add directly or indirectly to the profits of the enterprise.

A Simple Illustration

The best way to clarify the factors that enter into a rational investment decision is to take a simple example. In this example, we assume that the investment project is expected to last only one year and that at the end of the year the equipment is expected to be completely worn out so that it will have no scrap value. The assumption of a one-year life is not really necessary and will be dropped in the course of our discussion, but it has the advantage of permitting us to avoid temporarily the complications associated with compound interest.

We consider the case of a firm which is considering investment in new equipment costing $50,000 and, as just indicated, having an expected life of one year. If the firm is to make a rational decision in this matter, it will presumably prepare a calculation such as that shown in Table 8–1. Depending upon the size of the company and the sophistication of its management, such a calculation may range from a casual scrawl on the back of a discarded

---

[2] It may be noted that nearly all of the formal analysis contained in J. M. Keynes's *General Theory of Employment, Interest, and Money* (New York: Harcourt, Brace & Co., 1936), on which the present analysis rests rather heavily, neglects the effect of investment on productive capacity, assuming, in fact, that labor is the only variable factor of production. Indeed, Keynesian analysis is largely a study of the factors determining the extent of utilization of aggregate economic capacity, being in this respect somewhat analogous to the Marshallian short-run analysis of price theory, which studies the determinants of the rate of utilization of existing capacity in the case of a single firm or industry (see Alfred Marshall, *Principles of Economics* [8th ed.; New York: Macmillan Co., 1948], Book V, chap. v). However, as indicated above, Keynesian analysis, unlike Marshallian, suffers from an internal contradiction in the sense that it ordinarily involves net investment which, for the economy as a whole, must necessarily change economic capacity.

envelope to an extensive document, with supporting data based on estimates by engineers and accountants, prepared for submission to the officials charged with the responsibility for preparation of the firm's "capital budget."

Perhaps the most striking thing about the calculation shown in Table 8–1 is the number of times the word "expected" appears in it. Decisions concerning investment always involve quite complex forecasts concerning many variables relating to the position of the firm. It is necessary to estimate the expected additional revenue that will result from use of the new facilities

TABLE 8–1

Calculation of Expected Returns from a Hypothetical Investment Having Expected Life of One Year

Cost of the equipment		$ 50,000
Expected life of the equipment		1 year
Expected additional output resulting from the use of the equipment		10,000 units
Expected price per unit of output		$10
Expected additional revenue (10,000 units × $10)		$100,000
Expected additional expenses other than depreciation and interest		$ 47,000
Labor costs	$12,000	
Raw material costs	27,000	
Maintenance costs	2,000	
Property and excise taxes	4,000	
Miscellaneous costs	2,000	
Expected gross return on investment before depreciation and interest		$ 53,000
Expected depreciation cost		50,000
Expected return before deduction of interest cost		$ 3,000
Expected rate of return on investment before deduction of interest ($3,000 ÷ $50,000)		6%

and to set against these revenues the expected additional costs of labor, materials, depreciation, and so on that will be occasioned by the new equipment. In this particular example, the expected net monetary return amounts to $3,000 on an investment of $50,000, or a rate of return of 6 percent. We shall refer to the expected rate of return on an investment project, computed in this way, as the *marginal efficiency of investment* of the project.

The reader will note that interest on the funds invested in the equipment has not been included among the costs. Instead we shall handle interest by comparing the interest rate with the rate of return on investment in the equipment. Subject to some qualifications which are discussed below, in this example, for instance, we may suppose that if the interest rate that the firm must pay to borrow the funds to buy the machine is 5 percent, the firm will invest in the equipment, since it expects to earn $3,000 or 6 percent on the investment, whereas the cost of borrowing $50,000 for one year at 5 percent is only $2,500, leaving the firm with a net return of $500 after paying the interest. On the other hand, if it is necessary to pay an interest rate of 7 percent, the investment will not be made, since the interest cost will exceed the expected return on the investment. In general, the investment will be carried out if the marginal efficiency of investment exceeds the rate of interest.[3] Replacement investment as well as net investment should be governed by a comparison of the marginal efficiency of investment with the rate of interest. That is, in principle, reinvestment of depreciation allowances should not be carried out unless it will add to the profits of the stockholders.

Everyday discussion often makes a distinction between net investment designed to make possible an increase in output and investment designed to reduce the costs of producing the existing output. The latter may be termed *cost-cutting investment.* Actually, the distinction is not likely to be clear-cut, since in most cases investment which reduces the cost of producing the current output will also increase productive capacity. However, investment of a purely cost-cutting nature can be evaluated in a framework such as that shown in Table 8–1. In such cases, the increase in revenue will be zero, and the gain will come through a reduction in expenses. For example, the investment may make possible the replacement of expensive skilled labor with lower paid semiskilled or unskilled labor.

It is apparent from Table 8–1 that the calculation of the marginal efficiency of investment requires the firm to make judgments, either explicit or implicit, concerning a great number of variables. In this example, it is assumed that the cost of the equipment is known, but in many cases even this has to be estimated—the cost of building a plant or a store may be higher than was originally anticipated if the cost of labor or materials rises during the period of construction. In order to estimate the expected revenue from the investment, it is necessary to estimate (*a*) the physical productivity of the equipment, and (*b*) the price at which the product can be sold. If the firm has a monopoly of semimonopoly position, so that it is faced by a downward sloping demand curve for its product, it must estimate the extent to which the sale of additional units of product will depress the price it will receive on the sales it is already making.[4] Its estimate of the demand

---

[3] We are implicitly assuming that investment is financed by borrowing money at interest through the issuance of bonds or the flotation of loans. In practice, of course, business finances much of its expenditure for plant and equipment in other ways—notably through the use of so-called "internal funds" derived from retained profits and depreciation allowances and to a lesser extent through the issuance of new stock (equities). The existence of alternative sources of financing creates complications which will be taken up in Chapter 10.

[4] For example, suppose the firm is now selling one million units of its product at $10 apiece for total revenue of $10 million. If it installs new equipment which will permit it to produce 200,000 more units, it estimates that it will have to reduce the price to $9.50 in order to induce its customers to buy the additional units. In this case, the additional revenue resulting from installation of the equipment will not be $1.9 million (200,000 × $9.50). Rather it will be necessary to deduct $500,000 to allow for the fact that the price will have to be reduced by $0.50 per unit on the 10 million units now being sold. Thus the additional revenue resulting from the investment will be only $1.4 million.

conditions for its product may go astray for many reasons: poor judgment on its part in estimating the market for its product, the introduction of improvements in production technique or in the product itself on the part of the firm's competitors, an unexpected change in consumers' tastes for its product, or changes in demand associated with general business conditions.

Similarly with respect to cost estimates, the firm must estimate many things, including physical inputs of labor and materials per unit of output, wage rates, raw material prices, excise and property taxes at various levels of government, and so on. The equipment may not last as long as was originally anticipated, due either to wear and tear or to obsolescence, so that depreciation costs turn out to be higher than was expected.

In the example we are considering, it is apparent that a small miscalculation can easily wipe out the expected profit from the investment. For example, with the interest rate at 5 percent, if the price at which 10,000 units can be sold turns out to be only one half of 1 percent less than expected or if wage rates are 4 percent higher than anticipated, the investment will show no profit. Of course, mistakes may be made in either direction, and the profit may prove to be greater than was expected. However, it appears that most businessmen are averse to risk in the sense that they require a prospective profit in order to induce them to assume it. In this case, for example, the firm may not be willing to undertake the investment at an interest rate of 5 percent, because the expected return of $500 after deduction of interest cost is not sufficient to compensate for the risk assumed. In fact, depending upon the businessman's assessment of the risk involved and his attitude toward assuming risk, it is even possible that he would not undertake this investment if he were able to obtain the funds at no cost, since a net expected return of more than six percent might be required to compensate him for risk.

Generalization of the Marginal Efficiency Concept

There is another way of viewing the calculation shown in Table 8–1, which is for some purposes superior to that used in the table. The firm in this example invests $50,000 in the equipment at the beginning of the year and expects to receive back at the end of the year $53,000 after the deduction of all expenses except interest and depreciation. That is, the firm expects to get back its original investment plus $3,000 of profits. One way to compute the marginal efficiency of investment for this investment is to find the interest rate at which $50,000 would have to be invested in order to grow in value to $53,000 at the end of a year. If the marginal efficiency of investment is designated as i, we have the following equation

$$\$50{,}000(1 + i) = \$53{,}000$$

This equation could be written alternatively as

$$\$50{,}000 = \frac{\$53{,}000}{1 + i}$$

When the calculation is expressed in this way, i is said to be the rate at which the expected future gross returns (\$53,000) must be *discounted* in order to make them equal to the original cost. Solving for i, we obtain

$$1 + i = \frac{\$53{,}000}{\$50{,}000}$$

or

$$i = \frac{\$53{,}000}{\$50{,}000} - 1 = 1.06 - 1 = 0.06 \text{ or } 6 \text{ percent}$$

Thus, this method calculating the marginal efficiency of investment gives the same answer as that shown in Table 8–1.

When the expected life of the investment is more than one year, the problem of compounding interest makes it impossible to calculate the marginal efficiency of investment by the simple method used in Table 8–1. However, the discounting method can still be used. Suppose, for example, that an investment costs \$100,000, has an expected life of two years, and is expected to yield, after deduction of all expenses except depreciation and interest, returns of \$53,000 the first year and \$61,000 the second year. If we let X_1 stand for the amount of money which if invested now at a rate of return of i will accumulate to \$53,000 at the end of one year, we have

$$X_1(1 + i) = \$53{,}000$$

or

$$X_1 = \frac{\$53{,}000}{1 + i}$$

If X_2 stands for the sum of money which if invested now at a rate of i per year will accumulate to \$61,000 at the end of two years, we have

$$X_2(1 + i)^2 = \$61{,}000$$

or

$$X_2 = \frac{\$61{,}000}{(1 + i)^2}$$

If we set X_1 plus X_2 equal to \$100,000, the cost of the investment, and solve for i, we will obtain the marginal efficiency of investment. That is,

$$\$100{,}000 = \frac{\$53{,}000}{1 + i} + \frac{\$61{,}000}{(1 + i)^2}$$

or, multiplying through by $(1 + i)^2$, we have

$$\$100{,}000(1 + i)^2 - \$53{,}000(1 + i) - \$61{,}000 = 0$$

This is a quadratic equation in $(1 + i)$, the solution of which is

$$1 + i = 1.08975$$

or

$$i = 0.08975 \text{ or } 8.975 \text{ percent}$$

Thus, the *marginal efficiency of investment* in this case is 8.975 percent.

In accordance with this analysis, we may define the marginal efficiency of investment as the *rate of discount which must be applied to the expected returns from an investment project in order to make its present value exactly equal to its cost.*[5] "Returns" as used in this definition are equal to receipts minus all costs except depreciation and interest. Analytically, the definition may be presented as follows:

$$C = \frac{E_1}{1 + i} + \frac{E_2}{(1 + i)^2} + \frac{E_3}{(1 + i)^3} + \cdots + \frac{E_n}{(1 + i)^n} \qquad (1)$$

where C is the original cost of the project, E_1, E_2, E_3, . . . E_n are the expected gross (dollar) returns in each of the n years in the expected life of the project, and i is the marginal efficiency of investment. Given the cost of the investment, the expected returns in each of the years (the E's), and the expected life of the investment (n), it is possible to find the approximate value of i from this equation, although the process of estimation is generally awkward and difficult.

In the case in which the expected returns are the same in all years (i.e., $E_1 = E_2 = E_3 = \ldots E_n$), the expression given above can be simplified considerably and a solution can be obtained by the use of compound interest tables. If E is the expected yearly return, we have

$$C = \frac{E}{1 + i} + \frac{E}{(1 + i)^2} + \frac{E}{(1 + i)^3} + \cdots + \frac{E}{(1 + i)^n} \qquad (2)$$

[5] That is, the present value (V) of a payment of E dollars to be received, say, three years from now, with interest compounded once a year, when the interest rate at which funds can be invested is $100r$ percent is obtained by solving the equation

$$V(1 + r)^3 = E$$

for V—that is, by finding the sum which if invested at the specified interest rate will build up to E dollars at the end of three years. Solving explicitly for V, we have

$$V = \frac{E}{(1 + r)^3}$$

The present value of a series of payments (E's) to be received at specified future times is the sum of a series of terms like this, one for each payment. Thus, we get the expression in Equation 4.

Multiplying both sides of this equation by $1/(1+i)$, we obtain

$$\frac{1}{1+i}C = \frac{E}{(1+i)^2} + \frac{E}{(1+i)^3} + \frac{E}{(1+i)^4} + \cdots + \frac{E}{(1+i)^{n+1}} \qquad (3)$$

Subtracting Equation 3 from Equation 2 and canceling out terms, we have

$$\left(1 - \frac{1}{1+i}\right)C = \frac{E}{1+i} - \frac{E}{(1+i)^{n+1}}$$

Multiplying both sides by $(1+i)$ and rearranging terms, we have

$$C = \frac{E}{i}\left(1 - \frac{1}{(1+i)^n}\right)$$

By using this expression, it is possible to use compound interest tables to find i, given the values of C, E, and n. One special case is of some interest. If the investment project is indestructible—that is, can be expected to last virtually forever—the expected life (n) becomes infinite. For any (positive) value of i, $1/(1+i)^n$ approaches zero as n approaches infinity. Thus, for the indestructible project, we have simply

$$C = \frac{E}{i}$$

The student will doubtless recognize this as the familiar *capitalization formula.* This formula is often used to obtain an estimate of the market value of a piece of land (C), given the expected annual rent (E), and the capitalization rate (i), which represents the rate of return that the investor can obtain on other investments involving approximately the same degree of risk.

An Alternative Criterion

According to the above analysis, the firm that is interested in maximizing its profits will carry out all investments for which the marginal efficiency of investment is in excess of the interest rate it must pay for the funds and will reject all investment projects for which the rate of interest is in excess of the marginal efficiency of investment. There is another way of explaining this relationship which is useful for some purposes.

It was explained above that when the expected future returns from an investment are discounted at a rate equal to the marginal efficiency of investment, their sum is equal to the cost of the investment project. By a similar line of reasoning, when the expected returns are *discounted at the market rate of interest,* their sum is equal to the *present value* of the investment (see footnote 5). That is,

$$V = \frac{E_1}{1+r} + \frac{E_2}{(1+r)^2} + \frac{E_3}{(1+r)^3} + \cdots + \frac{E_n}{(1+r)^n} \qquad (4)$$

where V is the present value of the investment, the E's are the expected (dollar) returns, n is the expected life of the investment, and r is the market rate of interest. Comparing Equation 4 with Equation 1, it is apparent that if the rate of interest is equal to the marginal efficiency of investment ($r = i$), the present value of the investment is just equal to its cost ($V = C$). Moreover, since r appears with a positive sign in the denominators of all the terms on the right-hand side of Equation 4, it is clear that V rises when r falls and falls when r rises. That is to say, the higher the rate of interest at which the expected returns are discounted, the lower is the present value of an asset. From these relations, it is apparent that whenever the marginal efficiency of investment exceeds the rate of interest ($i > r$), the present value of the asset must exceed its cost ($V > C$); and that whenever the rate of interest exceeds the marginal efficiency of investment ($r > i$), the cost of the project must exceed its present value ($C > V$). The results of this discussion may be summarized as follows:

Interest Rate Criterion	*Present Value Criterion*	*Implied Action on Project*
$r < i$	$V > C$	Carried out
$r > i$	$V < C$	Rejected

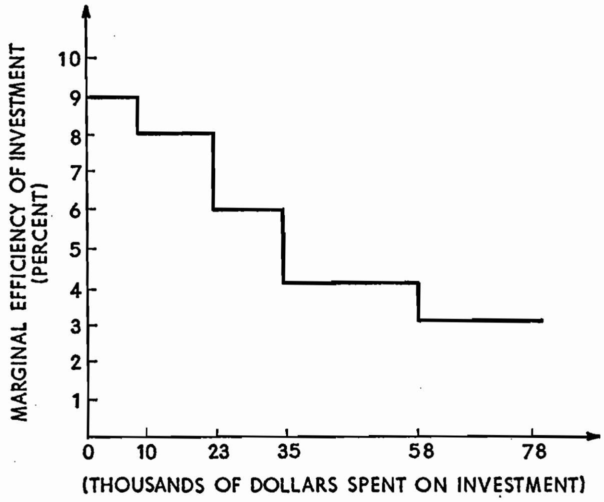
For some purposes, the present value criterion may be more useful than the interest rate criterion. For example, since the cost of investment will not be directly affected by a change in the rate of interest, we can judge the effects of interest rate changes on investment, to some extent at least, by considering the magnitude of the effects of such changes on the present value of the investment, since it is by raising present values above or depressing them below costs that changes in interest rates exert their effects on investment. Thus, in studying the sensitivity of investment to the rate of interest, it is often more illuminating to look at the problem in terms of changes in present values rather than in terms of changes in the interest rate directly.

The Marginal Efficiency of Investment Schedule

According to the above analysis, there is, in principle at least, a marginal efficiency of investment for any investment the firm might consider. If the projects under consideration are arranged according to their marginal efficiencies, it is possible to draw up a *marginal efficiency of investment schedule* for the firm. A hypothetical schedule of this kind is given in Figure 8–1. Let us suppose that the most attractive investment open to the particular firm we are considering is the investment of $10,000 in specialized equipment which would add to the efficiency of its plant. The marginal efficiency of this investment is 9 percent. Next in order of profitability is an investment of $13,000 in, let us say, office equipment to speed up the processing of data in the cost accounting section and to reduce the cost of these operations. This investment is estimated to have a marginal efficiency of 8 percent. Further investments in order of profitability are one of $12,000 having a marginal efficiency of 6 percent, one of $23,000 having a marginal efficiency of 4 percent, and one of $20,000 having a marginal efficiency of 3 percent.

FIGURE 8–1

Marginal Efficiency of Investment Schedule for a Single Firm

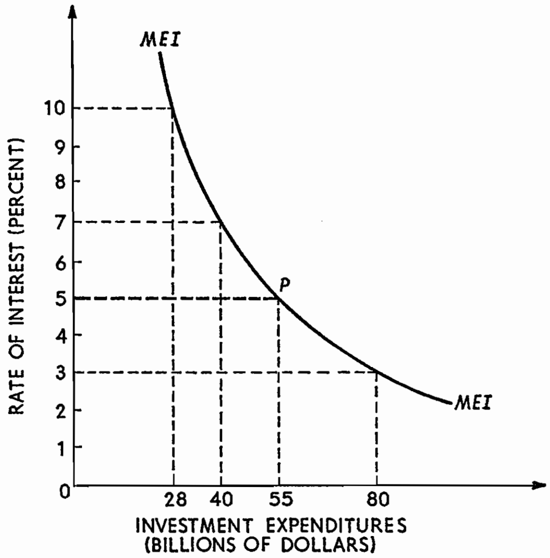


If the firm has this array of investments open to it, its marginal efficiency of investment schedule will be the broken line labeled *MEI* in Figure 8–1. If the interest rate it must pay to obtain funds is above 9 percent, it will invest nothing, since there are no investments that are prospectively profitable at this rate of interest. If the interest rate is between 8 and 9 percent, it will carry out only the $10,000 investment project having a marginal efficiency of 9 percent. If the interest rate is between 6 and 8 percent, it will carry out the $10,000 investment having a marginal efficiency of 9 percent and the $13,000 investment having a marginal efficiency of 8 percent, thus spending $23,000 in all. If the interest rate falls below 6 percent, the $12,000 investment having a marginal efficiency of 6 percent will be carried out, and so on. That is, for any interest rate, the broken line labeled *MEI* indicates the amount of investment spending that the firm will engage in. Thus, if it is assumed that decisions concerning investment are made by comparing the rate of interest with the marginal efficiency of investment, the marginal efficiency of investment schedule shows the amount of investment that will be carried out at each rate of interest.

It is useful to view the marginal efficiency of investment schedule for the economy as a whole, as illustrated in Figure 8–2, as a horizontal summation of the schedules of individual firms such as that shown in Figure 8–1. While the schedule of the individual firm is likely to be an irregular line, due to the "lumpiness" of investment projects, much of the irregularity is likely to disappear for the aggregative schedule, so that it is reasonably well depicted

FIGURE 8–2

Marginal Efficiency of Investment Schedule of the Nation



by a continuous smooth curve. Since the marginal efficiency of investment schedule indicates the amount of investment that will be carried out for each rate of interest, as explained above, the rate of interest is plotted on the vertical axis in Figure 8–2, and the *MEI* schedule may be viewed as the aggregative relation between the rate of interest and investment expenditures. Thus, the schedule indicates that at a rate of interest of 10 percent, investment would be $28 billion; and that as the interest rate fell, investment would increase, reaching, for example, $80 billion at a rate of 3 percent.

Several points should be noted about the marginal efficiency of investment schedule. First, investment is a *flow* rather than a *stock* concept—that is, it has a time dimension and indicates the amount of spending that would be done per unit of time, such as a year or a quarter. Second, the schedule is best viewed as an *instantaneous relationship* indicating different amounts of investment spending that would be initiated at alternative rates of interest during a specified period. Third, there are many *factors other than the rate of interest* that affect investment; these factors operate by changing the expectations underlying the investment calculations of firms, as explained above, and would be reflected in *shifts* of the *MEI* schedule to the right or to the left. These factors are discussed in the next chapter.

The Interest Elasticity of Investment

The interest elasticity of the marginal efficiency of investment schedule is a technical measure of the sensitivity of investment to changes in the rate of interest. This elasticity may be defined as

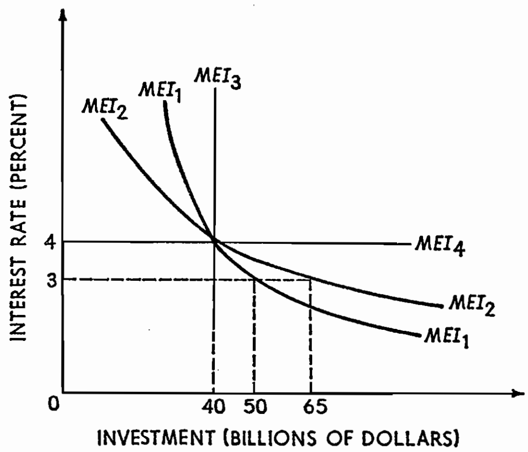
$$\frac{\text{Percentage change in investment expenditures}}{\text{Percentage change in the interest rate}} = \frac{\frac{\Delta I}{I}}{\frac{\Delta r}{r}} = \frac{r\Delta I}{I\Delta r}$$

Thus, if a change of 1 percent (of itself—e.g., from 3 to 3.03 percent) in the rate of interest causes a 2 percent change in investment, the elasticity is two; if a 1 percent change in the interest rate causes only a half of 1 percent change in investment, the elasticity is one half, and so on.

These ideas are illustrated in Figure 8–3. If the interest rate is 4 percent to start with and declines to 3 percent, investment will increase from $40 billion to $50 billion if the marginal efficiency of investment schedule is MEI_1, and from $40 billion to $65 billion if the schedule is MEI_2. Thus, the

FIGURE 8–3

Illustration of Alternative Degrees of Elasticity of Marginal Efficiency of Investment



schedule MEI_2 is more elastic than the schedule MEI_1 in this case.[6] The vertical line MEI_3 illustrates a completely inelastic schedule with investment entirely unaffected by changes in the rate of interest, while MEI_4 is a completely elastic schedule, which means that a slight decline in the interest rate below 4 percent will result in an indefinite increase in investment, while a slight rise in the rate above 4 percent will bring investment expenditures to a complete stop.

[6] In the above case, the interest elasticity of the curve MEI_1 appears to be one, since a 25 percent fall in the interest rate results in a 25 percent rise in investment. (Actually, it would be negative one, since a *fall* in the interest rate causes a *rise* in investment—i.e., the *MEI* schedule is negatively sloped). However, there is a difficulty arising from the customary practice in computing percentage changes of using the chronologically first value of the variable as the base. Thus, if the interest rate were to change in the other direction (from 3 to 4 percent), a 33⅓ percent rise in the interest rate would cause a 20 percent decline in investment and the elasticity would be one and two-thirds instead of one. To get around this problem and give the same estimate of elasticity whether the interest rate rises or falls, the concept of arc elasticity has been developed, which adopts the convention of using the average values of the interest rate and of investment between the two points in computing the elasticity whether the price rises or falls. On this basis, the elasticity would be

$$\frac{\dfrac{1 \text{ percent}}{3.5 \text{ percent}}}{\dfrac{\$10 \text{ billion}}{\$45 \text{ billion}}} = 1\tfrac{2}{7}$$

It should also be noted that because of the confusing but widespread practice in economics of measuring the independent variable on the vertical axis and the dependent variable on the horizontal axis (just opposite of the procedure ordinarily used in mathematics), a *steeply sloping* curve corresponds to a *low elasticity* and a *gently sloping* curve to a *high elasticity.*

Chapter 9

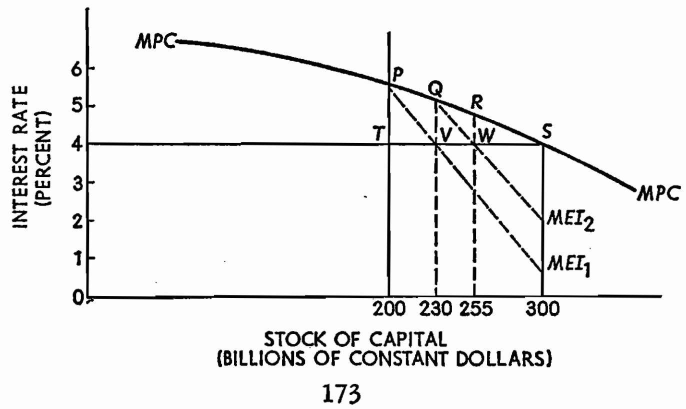
INVESTMENT II: THE DYNAMICS OF INVESTMENT

According to the analysis developed in Chapter 8, if the interest rate is given, the rate of investment spending will be determined at the point on the *MEI* schedule corresponding to that interest rate. For example, if the *MEI* schedule is that depicted in Figure 8–2 and the interest rate is 5 percent, investment expenditures will be $55 billion. The next question is: What are the factors that determine the position of the *MEI* schedule, and what causes it to change position from time to time?

RELATION BETWEEN INVESTMENT AND THE STOCK OF CAPITAL

Due to the operation of the law of diminishing returns, if the size and training of the labor force, the available technology, and the supply of natural resources are given, the expected incremental returns to additional units of capital will decline as the size of the stock increases. That is, there will be a marginal productivity of capital schedule such as the curve *MPC* in Figure 9–1, which relates the expected marginal return on capital to the stock of capital in existence. In a competitive economy, capital accumulation

FIGURE 9–1
The Marginal Productivity of Capital



will tend to be carried to the point at which the expected marginal productivity of capital is equal to the rate of interest—that is, to the point at which the expected returns from the last dollar of capital accumulated are equal to the cost of funds for investment. Thus, in Figure 9–1, if the rate of interest is 4 percent, the equilibrium stock of capital is $300 billion. Once this much capital has been accumulated, there will be no further growth of the capital stock unless the rate of interest falls below 4 percent or unless some development occurs which causes the marginal productivity of capital schedule to shift to the right. When $300 billion of capital has been accumulated, there will be just enough production of capital equipment to make good depreciation, but there will be no net investment.

The *marginal productivity of capital* should be distinguished from the *marginal efficiency of investment.* Capital is a *stock* of productive facilities existing at a particular time, while investment is that portion of the current *flow* of output which is devoted to the creation of additional productive facilities—thus, net investment is the change in the stock of capital during a specified period of time.

As an illustration of the difference between the two concepts, consider the situation depicted in Figure 9–1. Suppose that at the beginning of a certain year the stock of capital in existence is $200 billion so that we are at point *P* on the *MPC* curve. The interest rate is 4 percent, the corresponding equilibrium stock of capital being $300 billion corresponding to point *S* on the curve. Thus, there is an investment potential of $100 billion. However, when businessmen set out to take advantage of these investment opportunities, the heavy demand for capital goods presses on the limited productive capacity of the capital goods industries and drives up the prices of such goods. As indicated earlier, the marginal efficiency of investment is obtained by solving for i in Equation 1 in Chapter 8, which is repeated here for convenience.

$$C = \frac{E_1}{1+i} + \frac{E_2}{(1+i)^2} + \frac{E_3}{(1+i)^3} + \cdots + \frac{E_n}{(1+i)^n}$$

As the rise in capital goods prices increases the values of C for various investments, the corresponding values of i (the marginal efficiency of investment) are reduced, since with given expected returns (the E's), a rise in C requires a fall in i if the above equation is to continue to be satisfied. As a result, the marginal efficiency of investment schedule, instead of coinciding with the *MPC* curve between point *P* and point *S,* declines more rapidly than *MPC,* following a path such as the dashed curve MEI_1.[1]

The rate of investment for the first year is determined by the intersection

[1] J. M. Keynes, in *The General Theory of Employment, Interest, and Capital* (New York: Harcourt, Brace & Co., 1936), used the term "marginal efficiency of capital" to refer to the expected rate of return on new investment. We have changed the name of this concept to "marginal efficiency of investment" in order to avoid confusing it with the "marginal productivity of capital."

of the horizontal line at 4 percent interest and the curve MEI_1. Thus, net investment is $30 billion, and when this capacity is installed, the stock of capital is $230 billion, corresponding to point *Q* on the *MPC* curve. Accordingly the marginal efficiency of investment schedule for the second year is the curve MEI_2, which starts from point *Q* and declines more rapidly than the *MPC* curve due to the tendency for capital goods prices to rise as the rate of production of capital goods increases. If the interest rate remains at 4 percent, net investment for the second year will be $25 billion (the distance *VW*), and the stock of capital at the end of the second year will be $255 billion, corresponding to the point *R* on the *MPC* curve.[2]

The Dynamic Basis of Investment

It is clear that as long as the *MPC* curve remains in its original position, the rate of net investment will gradually decline until the capital stock has been built up to $300 billion, corresponding to point *S* on the *MPC* curve. Once this point has been reached, there will be no further net investment unless (1) the rate of interest is reduced below 4 percent, or (2) something happens to shift the *MPC* curve to the right. A continuing strong demand for capital which will sustain a substantial rate of net investment is dependent upon the existence of *dynamic forces which keep the MPC curve shifting more or less continuously to the right.* In the absence of such dynamic forces, net investment would eventually fall to zero. For a time it might be possible to stimulate investment by lowering the rate of interest, but there are likely to be limits to the possibility of encouraging investment in this way.

THE ACCELERATION PRINCIPLE

In terms of the analysis just presented, the determinants of investment may be classified into factors determining the stock of capital desired by businessmen and the speed with which they proceed to achieve that desired stock. Many writers have put considerable emphasis on the so-called "acceleration principle" as an explanation of investment behavior in this context.[3]

---

[2] The prices of capital goods assumed to prevail at successive points (such as *P, Q, R,* and *S*) on the *MPC* curve are those which would prevail when net investment was zero and the capital goods industries were producing at a rate just sufficient to make up for depreciation, thus keeping the capital stock constant at the specified level. The breakdown of the process of capital expansion into "years" is an oversimplification; both investment and growth of the capital stock are in fact continuous processes.

[3] An early exposition of the acceleration principle, which is still valuable, is to be found in J. M. Clark, "Business Acceleration and the Law of Demand: A Technical Factor in Economic Cycles," *Journal of Political Economy,* Vol. XXV, March 1917, pp. 217–35, reprinted with an additional note by the author in G. Haberler (ed.), *Readings in Business Cycle Theory* (Philadelphia: Blakiston Co., 1944), pp. 235–60. Excellent systematic discussions of the acceleration principle are to be found in Gottfried Haberler, *Prosperity and Depression* (3d ed.; Geneva: League of Nations, 1941), pp. 85–105; and R. A. Gordon, *Business Fluctuations* (New York: Harper & Bros., 1952), chap. 5.

Application to a Single Firm

Perhaps the best way to explain the acceleration theory of investment is to take a somewhat oversimplified hypothetical example. Suppose that in shoe manufacturing, technological or engineering requirements make it necessary to have $2 of capital invested in equipment for each dollar's worth of shoes produced and sold.[4] Suppose further that equipment has a uniform life of 20 years, that the particular shoe manufacturing firm we are considering is initially selling $100 million worth of shoes a year and has $200 million of capital invested, and that the age distribution of this capital is such that 1/20 or $10 million will wear out each year for the next 20 years.

The example is presented in Table 9–1. In this example, it is assumed

TABLE 9–1

Hypothetical Illustration of the Acceleration Principle in the Shoe Industry (millions of dollars)

Year	*Sales*	*Desired Stock of Capital*	*Actual Stock of Capital*	*Replacement Investment*	*Net Investment*	*Gross Investment*
1961	100	200	200	10	..	10
1962	110	220	220	10	20	30
1963	125	250	250	10	30	40
1964	135	270	270	10	20	30
1965	140	280	280	10	10	20
1966	140	280	280	10	..	10
1967	135	270	270	..	..	..
1968	125	250	260	..	..	..
1969	115	230	250	..	..	..
1970	100	200	240	..	..	..
1971	95	190	230	..	..	..
1972	95	190	220	..	..	..
1973	100	200	210	..	..	..
1974	115	230	230	10	20	30

that the behavior of sales is given; that is, the table does not explain the behavior of sales but rather explains the behavior of investment given the behavior of sales. In the initial year 1961, sales are $100 million and the stock of capital is $200 million. Thus, the stock of equipment is just sufficient to produce the output needed to satisfy the demand; accordingly the firm purchases $10 million of equipment to make up for depreciation, but does not add to its stock of equipment. Thus, replacement investment is $10 million, net investment is zero, and gross investment is $10 million.

[4] Throughout this example we assume that all prices are constant, so that dollar values can be thought to coincide with real values. That is, the relationship between output and investment capital is a physical one, but for the sake of convenience we express it in dollars.

The first point to notice in this example is that changes in the demand for shoes have a magnified effect on the demand for equipment to produce shoes. Thus, in 1962, it is assumed that sales of shoes rise to $110 million. In order to produce this amount of shoes it is necessary to have $220 million of equipment; accordingly, in addition to replacing the $10 million of machinery that wears out in that year, the firm purchases a further $20 million of equipment to build its stock of capital up to the required level. Thus, replacement investment is $10 million, net investment is $20 million, and gross investment is $30 million. Result: a 10 percent rise in sales of shoes from 1961 to 1962 produces a tripling—i.e., a 200 percent increase—of gross investment in equipment. It may be noted that the extent of magnification is greater the longer is the life of the equipment.[5]

The next point to notice is that investment in equipment depends on the *absolute rate of change of sales of the final product.* In 1963, sales rise by $15 million compared with the rise of $10 million in 1962, and investment in equipment rises to $40 million. In 1964, sales continue to increase but the absolute rate of increase declines from $15 million to $10 million, and this is enough to cause a decline in gross investment from $40 to $30 million. That is, a mere reduction in the *absolute rate of increase of sales causes a decline in investment.* The rate of increase of sales falls further, to $5 million in 1965, and investment drops again to $20 million. And in 1966, sales level off at $140 million, and this causes net investment to drop to zero, so that gross investment includes only the $10 million of replacement investment.

Since, under the acceleration principle investment drops as soon as the rate of increase of sales slows down, *turning points in investment precede turning points in sales.* That is, in our example investment reaches a peak of $40 million in 1963, whereas sales do not reach their peak of $140 million until two years later, in 1965.

Table 9–1 also brings out the fact that the effects of the acceleration principle are not symmetrical as between periods of rising final demand and periods of falling final demand. Beginning with 1966, there is a period of declining sales lasting until 1971. The adjustment of the stock of equipment to the declining rate of production is impeded, however, by the fact that the stock of equipment cannot decline faster than machines wear out. For example, between 1967 and 1968, sales drop by $10 million, but the stock of machines declines by only $10 million as a result of failure to replace units that wear out. Thus, the desired stock of equipment is only $250 million in 1968, while the actual stock of equipment is $260 million. This situation continues for several years, so that by 1971, there is "excess

[5] In our example, if everything were the same as in Table 9–1 except that the life of the equipment was reduced to 10 years, replacement demand would be $20 million per year and gross investment would increase from $20 million in 1961 to $40 million in 1962, a rise of 100 percent compared with the 200 percent rise shown in the table.

capacity" amounting to $40 million of capital equipment—that is, the desired stock of equipment is only $190 million, while the actual stock is $230 million. Consequently, even though sales level off at $95 million in 1972 and rise to $100 million in 1973, there is no investment at all in these years. Finally, in 1974, as a result of a sharp rise in sales to $115 million, together with the continued wearing out of existing equipment, there is a reappearance of investment demand amounting to $30 million in that year.

An Aggregative Accelerator Multiplier Model

A number of economists have constructed aggregative economic models built around the acceleration principle.[6] In the above example, the principle was applied to a single firm. To illustrate its application to the entire economy, consider a model with an output lag in which businessmen produce in period t an amount equal to their sales in period $t - 1$. Suppose further that they desire to have a dollars of capital for each dollar of sales made in the previous period and that they always engage in an amount of investment sufficient to achieve this objective. That is,

$$K_t = a\text{Sales}_{t-1} = aY_t \tag{1}$$

where K is the stock of capital and Y is gross national product.[7] The same relation holds for the previous period; thus

$$K_{t-1} = a\text{Sales}_{t-2} = aY_{t-1} \tag{2}$$

Subtracting Equation 2 from Equation 1, we have

$$K_t - K_{t-1} = aY_t - aY_{t-1}$$

Since investment is equal to the change in the stock of capital—that is, $I_t = K_t - K_{t-1}$, we have

$$I_t = a(Y_t - Y_{t-1})$$

Adding a consumption function and employing an output lag which makes the current period's output equal to the previous period's sales, we have the following model:

$$C_t = cY_t \tag{3}$$
$$I_t = a(Y_t - Y_{t-1}) \tag{4}$$
$$Y_t = C_{t-1} + I_{t-1} + A_{t-1} \tag{5}$$

[6] See P. A. Samuelson, "Interactions between the Multiplier Analysis and the Principle of Acceleration," *Review of Economic Statistics*, Vol. XXI, May 1939, pp. 75–78, reprinted in Haberler, *op. cit.*, pp. 261–69; also J. R. Hicks, *A Contribution to the Theory of the Trade Cycle* (3d ed.; Oxford, England: Clarendon Press, 1956).

[7] We assume, for present purposes, that capital goods have an infinitely long life, so there is no depreciation and no replacement. Thus, gross investment equals net investment and GNP equals net national product.

where C is induced consumption expenditures and A is autonomous expenditure (which might be investment expenditure not governed by the acceleration principle, an element of consumption independent of the level of income, or government purchases of goods and services). The constant a may be called the acceleration coefficient and the constant c is the marginal (and average) propensity to consume.

TABLE 9–2

Hypothetical Illustration of Accelerator Multiplier Model

Period	Y *Gross National Product* ($C_{t-1} + I_{t-1} + A_{t-1}$)	C *Induced Consumption* ($0.5Y_t$)	A *Autonomous Expenditure*	I *Induced Investment* ($Y_t - Y_{t-1}$)	S *Saving* ($Y_t - C_t$)
0	400.0	200.0	200.0	...	200.0
1	400.0	200.0	210.0	...	200.0
2	410.0	205.0	210.0	10.0	205.0
3	425.0	212.5	210.0	15.0	212.5
4	437.5	218.8	210.0	12.5	218.8
5	441.3	220.7	210.0	3.8	220.7
6	434.4	217.2	210.0	−6.8	217.2
7	420.4	210.2	210.0	−14.0	210.2
8	406.2	203.1	210.0	−14.2	203.1
9	398.9	199.4	210.0	−7.3	199.4
10	402.1	201.1	210.0	3.3	201.1
11	414.3	207.2	210.0	12.2	207.2
New equilibrium	420	210.0	210.0	...	210.0

NOTE: Details may not add to totals due to rounding.

In the numerical illustration shown in Table 9–2, the consumption and induced investment equations are as follows:

$$C_t = 0.5Y_t$$
$$I_t = Y_t - Y_{t-1}$$

Thus, the marginal propensity to consume (c) is 50 percent and the acceleration coefficient (a) is unity. We assume that autonomous expenditure is initially \$200 billion and that income is at the equilibrium level of \$400 billion, corresponding to that level of autonomous expenditures. The student can demonstrate for himself that if income is \$400 billion, consumption will be \$200 billion, which, together with \$200 billion of autonomous expenditure, provides just enough aggregate demand to absorb the \$400 billion of output. If income remains constant at this level, there will, of course, be no induced investment, since induced investment depends upon the *change* of income. This is the situation in period zero as shown in Table 9–2.

In period one it is assumed that autonomous expenditure (A) rises to

$210 billion, thus disturbing the existing equilibrium, and that autonomous expenditure remains at this higher level for an indefinite time. The successive entries in Table 9–2 show how income and the other variables will behave under these conditions. In period two income rises to $410 billion due to the initial impact of the rise in autonomous expenditure. This rise in income causes a rise of $5 billion in consumption in period two, since the marginal propensity to consume is 50 percent. Moreover the rise in income from $400 billion to $410 billion between periods one and two causes an additional $10 billion of induced investment in period two due to the operation of the acceleration principle. Thus, total sales of consumer goods and of investment goods rise to $425 billion in period two, thereby generating $425 billion of income in period three.

In this example, the existence of change-induced investment based on the acceleration principle introduces a self-generating cyclical fluctuation into the economy. Income continues to rise until it reaches $441.3 billion in period five, then declines until it reaches $398.8 billion in period nine, after which it turns up again. The cyclical fluctuation is the result of an interaction between induced consumption governed by the marginal propensity to consume and induced investment governed by the acceleration principle. The consumption function is a stabilizing factor, since the 50 percent marginal propensity to consume tends to slow down the changes in income. When such a slowdown takes place, it causes an absolute change in investment which is governed by the acceleration principle. For example, between periods two and three income rises by $15 billion, causing an equal amount of induced investment in period three. However, as a result of the fact that only half of the addition to income is spent on consumption, the rise in income is reduced to $12.5 billion between periods three and four, causing induced investment to drop to $12.5 billion in period four. Thus, the accelerator is thrown into reverse, and after one more period (in period five) the rise in income is brought to a stop and a decline sets in. However, after a few periods the interaction of the multiplier and accelerator checks the decline and causes income to rise again.[8]

In this case, the static multiplier is equal to the reciprocal of the marginal propensity to save, just as in the case of the simple multiplier analysis discussed in Chapters 6 and 7. That is, in this particular example the multiplier is 2, so that the rise in autonomous expenditure from $200 billion to $210 billion raises the equilibrium level of income by $20 billion from

---

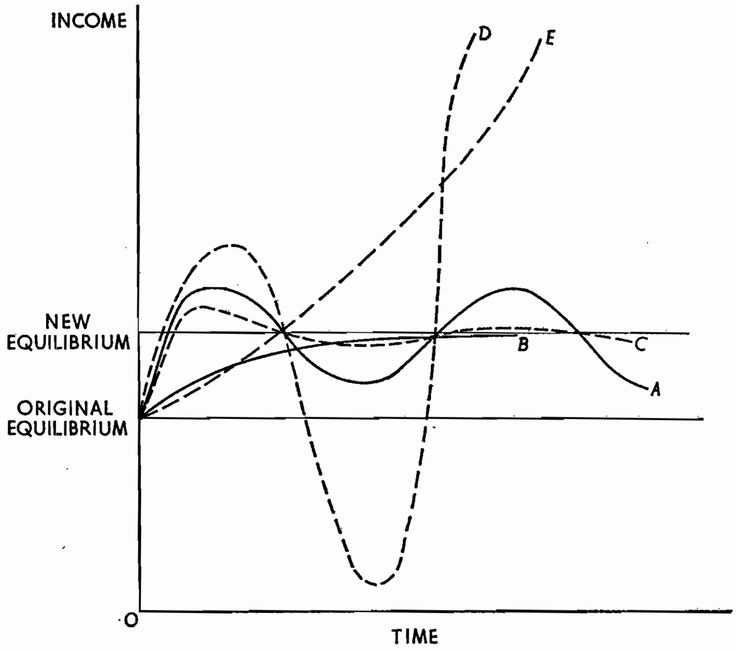
[8] In this example, no effort has been made to allow for the fact, referred to earlier and incorporated in the shoe industry example of Table 9–1, that induced investment cannot become negative in an amount greater than the rate at which equipment wears out—which, strictly speaking, in this example, would prevent any negative values of investment at all, since capital is assumed to have an infinitely long life. This refinement can be taken into account, as is done by Hicks, *op. cit.*

$400 billion to $420 billion. When income is $420 billion, consumption will be $210 billion (0.5 × 420), which, together with autonomous expenditure of $210 billion, produces aggregate demand of $420 billion, just sufficient to absorb the output being produced. As long as output remains at this level, there will, of course, be no induced investment.

However, the model illustrated in this example is *unstable.* When equi-

FIGURE 9–2

Alternative Time Paths of Income for Accelerator Multiplier Model



librium is disturbed by a rise in autonomous expenditure, as in our example, income begins to oscillate. In this particular case the oscillations are not damped, but will continue indefinitely with a constant amplitude unless a further shock impinges on the economy. Actually, this is only one of several possible patterns that may characterize the behavior of income in an accelerator multiplier model. The patterns of movement that might occur depend upon the magnitudes of the marginal propensity to consume (c) and the acceleration coefficient (a) and are illustrated in Figure 9–2. In addition to a constant amplitude fluctuation (Curve A), the other possibilities are (1) a steady approach toward the new equilibrium (Curve B); (2) damped oscillations (Curve C); (3) explosive oscillations (Curve D); and (4) steady and explosive geometric growth (Curve E).[9]

THE STOCK ADJUSTMENT MODEL

The accelerator multiplier model presented above has a number of shortcomings. Two of the more serious ones are the unrealistic assumption that any gap between the amount of capital desired by business and the amount actually in existence is always filled within a single period, and the failure to deal with depreciation and replacement. It is possible, however, to develop without great difficulty a more flexible version of the accelerator theory to get around these problems, at least to some extent.

If I is gross investment and D is depreciation, net investment will be equal to $I - D$. Since net investment in period t is, by definition, equal to the change in the stock of capital (K) between the end of period $t - 1$ and the end of period t, we have

$$I_t - D_t = K_t - K_{t-1} \tag{6}$$

Let us suppose—as is quite reasonable, at least as an approximation—that depreciation in period t is a constant proportion, p (the depreciation rate), of the capital stock in existence at the end of the previous period. That is

$$D_t = pK_{t-1} \tag{7}$$

We shall distinguish between the *actual* capital stock (K) and stock *desired* by businessmen (K^d). Let us suppose that the desired capital stock in period t is proportional to the volume of sales in the previous period. The factor of proportionality is a, which is closely analogous to the acceleration coefficient employed in the simpler model developed above. That is, $K_t^d = a\text{Sales}_{t-1}$. If production is adjusted to sales with a lag of one period, then $Y_t = \text{Sales}_{t-1}$. Thus, we have

$$K_t^d = aY_t \tag{8}$$

[9] By shifting Equations 3 and 4 back one period and substituting for C_{t-1} and I_{t-1} in Equation 5, the following autoregressive equation in income can easily be derived (treating A as a constant):

$$Y_t = (c + a)Y_{t-1} - aY_{t-2} + A$$

Given a value for A and values of Y for two initial periods, this equation can be used to trace out an unending sequence of values of Y through time. A *difference equation* such as this can be solved by analytical methods to obtain Y as a function of time, the parameters of the equation (c and a in this case), and certain initial conditions. This solution can then be analyzed to show how the time path of movement of Y is related to the values of the parameters. On difference equations and their solution, see W. J. Baumol, *Economic Dynamics: An Introduction* (2d ed.; New York: Macmillan Co., 1959), chaps. 9–11. For a more extenseive discussion of the possible patterns of behavior of the accelerator multiplier model used above, see A. J. Vandermeulen and D. C. Vandermeulen, *National Income: Analysis by Sector Accounts* (Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1956), pp. 297–302.

Finally, let us suppose that businessmen's investment decisions are designed to achieve the desired stock of capital but that they do not attempt to close the entire gap between actual and desired capital in a single period. Instead a certain proportion, s, of the gap is closed each period. That is, our theory of net investment is incorporated in the following equation

$$I_t - D_t = s(K_t^d - K_{t-1}) \tag{9}$$

For example, if the actual capital stock at the end of period $t - 1$ was 100, if the desired capital stock in period t (which, by Equation 8, depends on the sales of period $t - 1$ or the output of period t) is 110, and if s is 50 percent, investment in period t will be 5 $(0.50[110 - 100] = 5)$. The parameter s is often called the "speed of adjustment coefficient," since the greater its value, the more rapidly the gap between the desired capital stock and the actual capital stock is closed.

Since the left-hand sides of Equations 6 and 9 are the same, we can set the right-hand sides of the two equations equal to each other. This gives us

$$K_t - K_{t-1} = s(K_t^d - K_{t-1}) \tag{10}$$

Substituting the value of K_t^d from Equation 8 into Equation 10 and simplifying, we have

$$K_t = saY_t + (1 - s)K_{t-1} \tag{11}$$

If the value of s (the speed of adjustment coefficient) is one, it is apparent from Equation 10 that the entire gap between actual and desired capital is closed by net investment each period. Substituting $s = 1$ into Equation 11, we obtain

$$K_t = aY_t$$

The reader will note that this is the same as Equation 1 of the accelerator multiplier model developed above. Thus, the stock adjustment theory of net investment reduces to the simple accelerator theory in the special case in which $s = 1$.

Substituting pK_{t-1} for D_t (by Equation 7) in Equation 9 and rearranging terms, we obtain the following equation explaining gross investment:

$$I_t = s(K_t^d - K_{t-1}) + pK_{t-1} \tag{12}$$

We shall now develop a rather simple macroeconomic model incorporating this stock adjustment theory of investment. Let us suppose that personal consumption expenditures (C) are a constant proportion of disposable personal income (Y^d) with no lag. That is,

$$C_t = cY_t^d \tag{13}$$

In this model, the only item that must be deducted from GNP to obtain diposable personal income is capital consumption allowances (depreciation). Thus:

$$Y_t^d = Y_t - D_t$$

Substituting pK_{t-1} for D_t (by Equation 7), this becomes

$$Y_t^d = Y_t - pK_{t-1}$$

and upon substitution into Equation 13, we have

$$C_t = cY_t - cpK_{t-1} \tag{14}$$

Also, substituting pK_{t-1} for D_t in Equation 6 and rearranging terms, we obtain

$$K_t = K_{t-1} + I_t - pK_{t-1} \tag{15}$$

We are assuming an output lag, which gives us an equation setting output in period t equal to sales in period $t - 1$. That is,

$$Y_t = C_{t-1} + I_{t-1} + A_{t-1} \tag{16}$$

where A is autonomous noninvestment expenditure, which might include government purchases of goods and services or an element of consumption expenditures that is independent of income.[10]

It will be useful at this point to collect together the equations that form the model. They are

$$K_t^d = aY_t \tag{8}$$
$$I_t = s(K_t^d - K_{t-1}) + pK_{t-1} \tag{12}$$
$$C_t = cY_t - cpK_{t-1} \tag{14}$$
$$K_t = K_{t-1} + I_t - pK_{t-1} \tag{15}$$
$$Y_t = C_{t-1} + I_{t-1} + A_{t-1} \tag{16}$$

We leave it as an exercise for the student to show that, if the actual capital stock is fully adjusted to the desired capital stock each period and if there is no depreciation (that is, if $s = 1$ and $p = 0$), this model reduces to the simple accelerator multiplier model given by Equations 3, 4, and 5.

We shall assume that autnomous expenditure, A, takes on a constant value that does not change from period to period. For any specified value of A, there are equilibrium values of Y and K, which can be derived in the following way. First, substituting from Equation 8 into Equation 12, and then from Equations 12 and 14 into Equation 16, we obtain

$$Y_t = cY_{t-1} - cpK_{t-2} + s(aY_{t-1} - K_{t-2}) + pK_{t-2} + A \tag{17}$$

Next, substituting from Equation 8 into Equation 12 and then substituting from Equation 12 into Equation 15 yields

[10] In this case, complications arise if there are autonomous *investment* expenditures that are not governed by the stock adjustment mechanism, because these expenditures add to the stock of capital, thereby affecting depreciation and replacement requirements. These complications can be handled, but they make the analysis considerably more complex. Hence, we assume that all investment is induced by the stock adjustment mechanism.

$$K_t = K_{t-1} + s(aY_t - K_{t-1}) \qquad (18)$$

Now, when the economy is in equilibrium, income and the stock of capital will remain constant from period to period. If we let Y represent equilibrium income and K represent the equilibrium capital stock, we can assume that $Y = Y_t = Y_{t-1}$ and $K = K_t = K_{t-1} = K_{t-2}$. Substituting Y and K throughout Equations 17 and 18 in accordance with this assumption, we obtain

$$Y = cY - cpK + s(aY - K) + pK + A \qquad (19)$$

$$K = K + s(aY - K) \qquad (20)$$

From Equation 20, we have

$$K = aY \qquad (21)$$

which merely says that in equilibrium the actual capital stock bears the desired relation (given by Equation 8) to income. Substituting this into Equation 19, we have

$$Y = cY - cpaY + s(aY - aY) + paY + A$$

Solving this equation for Y, we obtain

$$Y = \frac{1}{1 - c(1 - ap) - ap} A \qquad (22)$$

Since, by Equation 21, $K = aY$, we also have

$$K = \frac{a}{1 - c(1 - ap) - ap} A \qquad (23)$$

Equations 22 and 23 give the equilibrium values of Y and K for specified values of autonomous expenditure, A, and of the parameters of the model, c, p, and a. The student will note that the speed of adjustment coefficient, s, does not affect the equilibrium solutions for Y and A, although as we shall see, it does affect the dynamic process of getting from one equilibrium position to another.

Translating Equation 22 into terms of differences in Y and A, we obtain the static multiplier applicable to a change in autonomous expenditure,

$$\frac{\Delta Y}{\Delta A} = \frac{1}{1 - c(1 - ap) - ap} \qquad (24)$$

It is necessary to explain the appearance of the term ap twice in the denominator of this multiplier. When income is in equilibrium, the stock of capital will remain constant and will be equal to aY. Depreciation will therefore be apY, net investment will be zero, and gross investment will be apY, just sufficient to make up for depreciation and hold the capital stock constant. Depreciation of apY must be deducted from Y to obtain disposable personal income; thus the marginal propensity to consume out of GNP is $c(1 - ap)$. In addition, for any rise in equilibrium income, ΔY, there will be an increase in the equilibrium capital stock of $a\Delta Y$, and this will require an increase of $ap\Delta Y$ in replacement investment each period to offset depreciation on the additional capital. The additional induced investment will make the static multiplier larger, just as does induced consumption, and this accounts for the last term, $-ap$, in the denominator of the multiplier.

In order to understand the stock adjustment model and see how it differs from the simpler accelerator multiplier model, it will be helpful to consider a numerical example. To make the example as similar as possible to that for the accelerator-multiplier model as shown in Table 9–2, we will assume that the marginal propensity to consume out of disposable personal income (c) is 50 percent; the desired capital-output ratio (a), which is analogous to the acceleration coefficient in the multiplier accelerator model, is unity; and autonomous expenditures (A) are \$200 billion. These are the same values that were used in the earlier example. In addition, we will assume that 25 percent of any gap between the desired and the actual capital stock is filled in each period, so that the speed of adjustment coefficient (s) is 25 percent; and that the life of capital goods is 50 periods, giving a depreciation rate (p) of 2 percent. To sum up, we assume the following values: $A = \$200$ billion, $c = 0.5$, $a = 1$, $s = 0.25$, and $p = 0.02$.

Substituting these values into Equation 22 to obtain the equilibrium value of Y, we have[11]

$$Y = \frac{1}{1 - 0.5(1 - (1)(0.02)) - (1)(0.02)}(200)$$

$$Y = \$408.2 \text{ billion}$$

Since a is unity, the desired capital stock (K^d) is also \$408.2 billion and since the system is in equilibrium the actual capital stock (K) is the same as the desired capital stock. Depreciation at a rate of 2 percent on this capital is \$8.2 billion; deduction of this amount from GNP of \$408.2 billion leaves disposable personal income of \$400 billion. Consumption is 50 percent of disposable personal income, or \$200 billion. Gross investment expenditures amount to \$8.2 billion, enough to offset depreciation and hold the capital stock constant. These amounts are entered in the first ("initial equilibrium") column of Table 9–3.

In the next period (period one in Table 9–3) GNP is equal to the sum of expenditures—consumption (\$200 billion), gross investment (\$8.2 billion), and autonomous expenditure (\$200 billion)—in the initial equilibrium period; that is, it is \$408.2 billion, the same as in initial equilibrium. Therefore, all the other entries for period one are the same as in initial

[11] In all of the calculations both in the text and in Table 9–3, the calculated values have been rounded off to the nearest 10th of a billion dollars.

TABLE 9-3

Hypothetical Illustration of Stock Adjustment Model*
(billions of dollars)

Symbol	Definition	How Calculated**	Number of Equation in Text	Initial Equilibrium	Time Period (t)								New Equilibrium
					1	2	3	4	5	6	7	8	
Y_t	GNP	$C_{t-1} + I_{t-1} + A_{t-1}$	16	408.2	408.2	418.2	425.7	430.7	433.6	434.7	434.7	433.9	428.6
C_t	Consumption expenditures	$cY_t - cpK_{t-1}$	14	200.0	200.0	205.0	208.7	211.2	212.6	213.1	213.1	212.7	210.0
I_t	Investment expenditures	$s(K_t^d - K_{t-1}) + pK_{t-1}$	12	8.2	8.2	10.7	12.0	12.4	12.1	11.6	10.8	10.1	8.6
A	Autonomous expenditures	Autonomous	..	200.0	210.0	210.0	210.0	210.0	210.0	210.0	210.0	210.0	210.0
K_t^d	Desired capital stock	aY_t	8	408.2	408.2	418.2	425.7	430.7	433.6	434.7	434.7	433.9	428.6
K_t	Actual capital stock	$K_{t-1} + I_t - pK_{t-1}$	5	408.2	408.2	410.7	414.4	418.5	421.3	425.9	427.7	429.3	428.6

* Values are rounded off to the nearest 10th of a billion.

** Parameter values assumed: $c = 0.5$, $a = 1$, $s = 0.25$, $p = 0.02$.

equilibrium, except autonomous expenditure, which is assumed to rise by \$10 billion to a new level of \$210 billion starting in this period. This raises GNP to \$418.2 billion in period two, also raising the desired capital stock to \$418.2 billion. Depreciation on the actual capital stock of period one is \$8.2 billion. Deducting this amount from the \$418.2 billion of GNP of period two leaves disposable personal income of \$410 billion. Consumption is 50 percent of this amount, or \$205 billion. Investment in period two is \$10.7 billion; this includes \$8.2 billion for replacement and net investment of \$2.5 billion, which represents 25 percent of the difference between the desired capital stock in period two (\$418.2 billion) and the actual capital stock of period one (\$408.2 billion). It is suggested that the student work out for himself the entries in Table 9–3 for a few more periods in order to strengthen his understanding of the model.

The static multiplier is given by Equation 24. Substituting the assumed values of c, a, and p, we have

$$\frac{\Delta Y}{\Delta A} = \frac{1}{1 - 0.5(1 - (1)(0.02)) - (1)(0.02)} = 2.04$$

Thus, the increase of \$10 billion raises *equilibrium* income by \$20.4 billion to a level of \$428.6 billion. The new equilibrium position of the economy is shown in the last column of Table 9–3.

As Table 9–3 shows, the increase of \$10 billion in autonomous expenditure starts a cyclical movement in economic activity. GNP reaches a peak of \$434.7 billion in period six and then begins to decline.[12] If the calculations were carried out for a larger number of periods, they would show that the decline will eventually end and GNP will turn up again. The pattern is different, however, from that shown in Table 9–2 for the accelerator multiplier example. In that case, the oscillations continued indefinitely with a constant amplitude, following a path similar to Curve A in Figure 9–2. Thus, that model is unstable in the sense that GNP fluctuates indefinitely around its equilibrium value without approaching it. In the stock adjustment example of Table 9–3, however, the amplitude of the oscillations gets smaller as time passes and GNP moves more and more closely to its equilibrium value in the course of its fluctuations, following a *damped* path such as Curve C in Figure 9–2. In other words, this model is stable in the sense that GNP approaches its equilibrium value in the course of time.[13]

[12] Table 9–3 shows GNP of \$434.7 billion in both period six and period seven. If the calculations were carried to a further decimal place, however, they would show that GNP in period seven is slightly below that in period six.

[13] The stock-adjustment model can be analyzed in the same way as the accelerator multiplier model as explained briefly in footnote 9, although the analysis is somewhat more complicated because the stock adjustment model contains more parameters. The difference equation in income for the stock adjustment model embodied in equations 8, 12, 14, 15 and 16 is as follows:

Another indication of the greater stability of the stock adjustment model lies in a comparison of the amplitude of movement of GNP during the first few periods. In Table 9–2, an increase of $10 billion in autonomous expenditure causes GNP to move up from its initial value of $400 billion in period one to a peak of $441.3 billion in period five; thus the multiplier for the period from the starting point to the initial peak is 4.13 (41.3 ÷ 10). In Table 9–3, on the other hand, a $10 billion increase in autonomous expenditure causes GNP to rise from $408.2 billion in period one to a peak of $434.7 billion in period six, for a multiplier of 2.65 (26.5 ÷ 10).

In models of this kind, the accelerator or stock adjustment mechanism is a destabilizing influence, while the fact that consumption increases by only a fraction (in these cases, one half) of income is a stabilizing influence. In the accelerator multiplier example of Table 9–2, these forces are in exact balance, so that fluctuations of constant amplitude are set in motion by a disturbance such as a change in autonomous expenditure. In the stock adjustment example of Table 9–3, the stabilizing influence of fractional consumption is dominant, with the result that the fluctuations are damped. Since the desired capital-output ratio is unity in both examples, it is apparent that the damping in the stock adjustment case is due to the fact that only a fraction (in this case, one fourth) of any gap between actual and desired capital is filled in each period.

CONCLUSION CONCERNING THE ACCELERATION PRINCIPLE

In its simplest form the acceleration principle is based on the naive assumption that a certain amount of capital is technologically required to produce a unit of output. This is not a tenable assumption, however, since it is possible to expand output from a given plant substantially, at least in the short run, by such devices as overtime work, multiple shift operations, the use of obsolete equipment, and so on. In any case, such a rigid technological explanation can scarcely be used to support the stock adjustment version of the acceleration principle, since under this version the stock of capital is only partially adjusted to the level of sales in a single period.

A more tenable but much less rigid basis for a somewhat modified form of the acceleration principle is that businessmen regard the current level of sales as an indicator of future demand for their product and make their investment decisions with a view to having the optimum stock of capital available at the time when that demand materializes. The optimum stock of

---

$$Y_t = (1 - s(1 - a) + c)Y_{t-1} - (c(1 - s) + sa(1 - p(1 - c)))Y_{t-2} + sA$$

(There is a similar difference equation for the capital stock, K, which can be derived without great difficulty and used to derive the above equation for income.) It may be noted that if $s = 1$ and $p = 0$, this reduces to the difference equation for the accelerator multiplier model given in footnote 9.

capital for any level of demand is the stock that would be expected to maximize the profits derived from meeting that demand. Since alternative methods of production are often available which use different quantities of capital relative to labor inputs, the optimum stock of capital for a given level of product demand will depend on the relative prices of the factors of production and on the rate of interest. For example, if the price of capital goods is low relative to the wages of labor, it will be profitable to adopt capital-intensive methods of production which enable the producer to use less labor or to substitute low-wage unskilled machine operators for high-wage skilled workers. Similarly, a low rate of interest will encourage the producer to use more capital in meeting a given level of demand.

If, for simplicity, we treat factor prices as given and unchanging, the quantity of capital desired to meet any level of demand for the product will depend inversely on the rate of interest. Thus, if the current level of income (equal to the previous period's volume of sales) is an index of future product demand to which producers attempt to adjust the stock of capital by investment activity, and if the amount of capital that is optimal for any level of demand depends on the rate of interest, we may regard the acceleration principle as operating in a somewhat muted way to shift the marginal efficiency of investment schedule back and forth in response to changes in current production and income. The operation of the acceleration principle is muted partly by the fact that only a portion of the gap between the actual and desired stock of capital is filled in a single period, according to some approximation to the stock-adjustment mechanism explained above. In addition, the relation is rendered somewhat inexact because businessmen surely do not regard every random aberration of their sales as an indication of a permanent change in future demand that requires an adjustment in the stock of capital.

Thus, the acceleration principle quite clearly plays a role in determining investment expenditures and contributes an element of inherent self-generating cyclical instability to economic activity, although it does not operate in any such mechanical way as the models developed above might suggest, instructive as these models are of the fundamental idea. It should be pointed out that the acceleration principle can also be applied to the demand for consumer durable goods. If, for example, the quantity of *services* of automobiles that consumers desire depends on the level of income, and if there is an optimum stock of automobiles needed to provide these services, the stock of automobiles will depend on the level of income, and net purchases of automobiles (the change in the stock) will depend on the *change* in income. There can be little doubt that it would be preferable to treat the demand for consumer durable goods in this way rather than simply to regard current purchases of these goods as a part of consumption that is primarily determined by income, as was done in Chapter 5 of this book, although the latter approach yields reasonably satisfactory results.

AUTONOMOUS FORCES AFFECTING INVESTMENT

Although the acceleration principle certainly plays an important role in determining investment, it would surely be a mistake to regard it as the only force at work. There are dynamic forces, classified below under the headings of population growth and technological progress, which are largely independent of current income, which exert important effects on the demand for capital and therefore on investment.

Population Growth

One dynamic factor which tends to shift the marginal productivity of capital curve to the right, thus sustaining the demand for capital goods, is the prospective growth of population. An increase in the expected future population will tend to increase the demand for capital (*a*) by increasing expected future levels of demand for consumer goods, and (*b*) by increasing the expected future labor force to provide more labor to work with any given amount of capital, thereby increasing its productivity.

Technological Progress

Technological advances, such as the development of new products or methods of production, generally help to sustain the demand for capital and to shift the *MPC* curve to the right. This is because the implementation of most technical changes involves the construction of new plants or the installation of new equipment. Particularly important in sustaining investment is the growth and development of new major industries. Thus, the spread of a railroad network across the United States was a major source of the demand for capital in the last half of the 19th century, the growth of the automobile industry was a major factor after the first World War, and the developments in electronics, including radio and television, have been especially important in more recent times.

The Long-Run Adequacy of Investment Demand

The above discussion suggests that a country which has a population which is growing rapidly and—more important—is expected to continue growing rapidly in the future, and in which the pace of technological development is rapid will have, on the average from year to year, a continuing heavy demand for new productive facilities. On the other hand, a country with a rather static population and a slow pace of technical development may find that investment demand is chronically rather weak.

During the Great Depression of the 1930's, the level of investment in the

United States was generally inadequate to provide outlets for the saving that would have occurred at a full-employment level of national income. It was this fact to which many economists attributed the deep and long-continued depression of that decade, which was not fully conquered until the early 1940's when the increase in government expenditures for armament associated with World War II produced the necessary stimulus. In fact, a school of thought developed among economists during the 1930's which contended that the depressed level of economic conditions of that period was more than merely an ordinary cyclical depression. It was said that the persistent weakness in the economy was a product of the slowdown in the rate of population growth which had been occurring for some years and which was expected to continue in the years ahead, together with a reduced pace of technological advance resulting particularly from the failure of great new industries to make their appearance to replace the demand for capital that had been generated by the growth of the automobile industry to maturity during the 1920's. This view, known as the *stagnation thesis,* attracted many adherents and led to the advocacy of a more active role on the part of the government in maintaining the level of aggregate demand through heavy public works expenditures and continuing deficit financing.[14]

The views of the so-called "stagnationists" were hotly disputed by other economists,[15] but before the issue was settled—if indeed it ever could have been settled—it was rendered academic by the surge of wartime military spending which not only restored the economy to full employment but went further and generated troublesome inflationary pressures. Undoubtedly the stagnationists underrated the ability of flexible prices and interest rates to facilitate adjustment of the economy to secular changes in the level of investment associated with population and technological changes. At the same time, it is quite conceivable that under some circumstances a continuing weakness in the demand for capital might create a chronic problem of achieving a sufficiently high level of aggregate demand to maintain full employment. On the other hand, it is also possible that under different circumstances, an excessively rapid rate of population growth and an unduly accelerated pace of technological advance might so stimulate the demand for

---

[14] For a clear statement of the "stagnationist" view, see A. H. Hansen, "Economic Progress and Declining Population Growth," *American Economic Review,* Vol. XXIX, March 1939, reprinted with minor revision in Haberler, *op. cit.,* pp. 366–84. See also Hansen's *Fiscal Policy and Business Cycles* (New York: W. W. Norton & Co., Inc., 1941). For a postwar study, stressing different causal factors but coming to somewhat similar conclusions, see Joseph Steindl, *Maturity and Stagnation in American Capitalism* (Oxford, England: Basil Blackwell, 1952).

[15] See, for example, G. W. Terborgh, *The Bogey of Economic Maturity* (Chicago: Machinery and Allied Products Institute, 1945). The issues are further discussed in A. H. Hansen, "Some Notes on Terborgh's 'The Bogey of Economic Maturity' " and D. M. Wright, " 'The Great Guessing Game': Terborgh versus Hansen," *Review of Economic Statistics,* Vol. XXVIII, February 1946, pp. 13–17 and 18–22, respectively.

capital as to raise a continuing problem of keeping the level of aggregate demand within the limits imposed by the supply of available resources; if this should happen, we might be faced with the problem of dealing with *secular inflation.* Which of these problems—if either—is to confront us in the years ahead is essentially unpredictable.

Chapter 10

INVESTMENT III: SOURCES AND COSTS OF INVESTMENT FINANCING

In the discussion of the marginal efficiency of investment schedule in Chapter 8, the schedule of the supply of funds available for financing investment was formulated in a rather naïve and unsophisticated way. In Figure 8–2 the curve *MEI* is the investment demand schedule, and it was stated that investment would be carried to the point at which the interest rate was equal to the marginal efficiency of investment. The implication of this statement appeared to be that all investment was financed by borrowing at the market interest rate. If, for instance, the interest rate was 5 percent, investment would be determined at point *P* on the *MEI* schedule. Investment expenditures would in this case occur at a rate of $55 billion per year. This analysis was, however, greatly oversimplified in its implicit assumption that all investment was financed by borrowing. In fact, investment is financed primarily by funds derived from three different sources: (1) gross business saving—so-called internal funds—which includes retained earnings and capital consumption allowances, (2) borrowing by means of loans and the issuance of new corporate bonds, and (3) the issuance of new corporate stock.

Table 10–1 shows the capital expenditures of the corporate nonfinancial business sector of the economy in 1968, together with the sources of funds applied to the financing of these expenditures.[1] Of the $79 billion spent on the accumulation of fixed capital and inventories, $63.1 billion, or about 80 percent, was covered by funds derived from retained profits and capital consumption allowances. This left $15.9 billion to be raised from external sources. However, business firms need to hold money and other financial assets, such as securities which they can sell to raise funds if necessary, to meet financial commitments as they arise and as a reserve against emergencies; and the amount of these holdings needs to increase from year to year as the volume of business expands. In 1968, corporations raised $46.5 billion from external sources, of which they used $30.5 billion to add to financial

[1] In recent years, corporations have accounted for roughly two thirds to four fifths of the total capital expenditures of nonfinancial business. In 1968, the ratio was 79.4 percent.

assets and the remaining $15.9 billion to cover the gap between their capital expenditures and the amount of funds derived from internal sources.[2]

Changes occur from year to year in the sources from which the funds for financing investment are derived, and 1968 should not necessarily be regarded as a typical year. However, Table 10–1 does give a general idea of the situation, and it is possible to make several crude generalizations about financing.

1. Internal sources are consistently dominant, accounting in recent years for from about 80 to over 100 percent of the total amount of capital expenditures. Retained earnings have increased as corporate profits have expanded with the growth of the economy, and depreciation allowances (which are the main component of capital consumption) have increased rapidly in line with the postwar growth of the stock of capital.

TABLE 10–1

Capital Expenditures and Financing, Corporate Nonfinancial Business, 1968
(amounts in billions)

(1) Capital expenditures		$79.0
Fixed investment	$72.5	
Change in inventories	6.5	
(2) Internal funds (= Gross corporate saving)		63.1
Retained corporate earnings	18.8	
Capital consumption allowances	44.3	
(3) Net external financing (1) − (2) = (4) − (5)		15.9
(4) External funds raised		46.5
From loans	14.1	
From bond issues	12.9	
From stock issues	−0.8	
From increase in other liabilities	20.4	
(5) Net acquisition of financial assets*		30.5
Deposits and currency	3.7	
Securities and other	26.8	

* Includes statistical discrepancy.
NOTE: Items may not add to totals due to rounding.
SOURCE: Federal Reserve System.

2. The relation between capital expenditures and internal funds varies with business conditions. During periods of economic expansion, capital expenditures rise rapidly due to the operation of acceleration effects described in Chapter 9, and retained earnings expand as the rise in GNP brings an increase in corporate profits. However, in such periods the rise in capital expenditures typically outstrips the rise in internal funds, with the result that increasing reliance must be placed on funds derived from external sources.

[2] This presentation has been somewhat oversimplified by the inclusion of the statistical discrepancy between the sources and uses of corporate funds in the figure for net acquisition of financial assets. This was done in the interest of simplifying the presentation and does not seriously distort the picture.

Thus, in recession years such as 1949, 1954, and 1958, internal funds have been more than sufficient to cover capital expenditures. On the other hand, during the long period of economic expansion from 1961 to 1968, capital expenditures rose from $37 billion to $79 billion, an increase of $42 billion, while internal funds expanded from $35.6 billion to $63.1 billion, a rise of $27.5 billion. Thus, the volume of funds that had to be raised from external sources increased steadily and dramatically during this period.

3. The sale of new stock is of almost negligible importance as a source of external funds. Nearly all external financing takes the form of loans from banks or other sources, or of new bond issues. There are several reasons for the primary reliance on debt financing. First, interest on loans or bonds is a deductible expense in calculating income under the corporate income tax, whereas dividends paid on stock are not deductible. With combined federal and state corporate income tax rates of around 50 percent, the deductibility of interest sharply biases financing decisions toward issuance of debt rather than equities. In addition, the financial institutions, notably life insurance companies, that have traditionally been the main suppliers of long-term funds to corporations are much more disposed to invest in bonds, which pay a fixed income, than in stocks. And, finally, sale of new voting stock to outsiders may weaken the position of the group currently controlling the corporation. It is important to note that retention of earnings adds to the equity capital of the corporation, but, unlike the sale of new stock to outsiders, leaves the control of the corporation undisturbed.

4. In deciding how to finance its capital expenditures, a corporation will be concerned about the maintenance of an appropriate balance among the elements of its capital structure. Debt financing adds an element of fixed costs which will increase the instability of the stockholders' return on investment. For example, suppose a corporation has a total capitalization of $125 million, of which $100 million is equity capital and the remaining $25 million is debt bearing interest at 5 percent. In a period of prosperity when the return on its total investment is 8 percent or $10 million, the corporation will have to pay $1.25 million in interest, leaving $8.75 million of profits for stockholders. Thus, the return on equity is 8.75 percent instead of the 8 percent return that would be realized if the entire $125 million was equity capital. On the other hand, if in a recession the total return on capital fell to 3 percent or $3.75 million, the return to stockholders would fall to $2.5 million or 2.5 percent instead of the 3 percent that would be realized if the entire $125 million was equity capital. As the fraction of debt in the total capitalization increases, the degree of cyclical variability in the return to stockholders becomes progressively greater. In addition, the fixed charges associated with a large component of debt in the capital structure create a danger of financial difficulties which may in an extreme case lead to bankruptcy.

THE SUPPLY-OF-FUNDS SCHEDULE

Let us attempt to construct a supply-of-funds schedule for a firm on the assumption that there are three sources from which it may obtain funds to finance investment: (1) retained earnings and depreciation allowances, (2) borrowing, and (3) issuance of new equities.[3] The shape of the supply schedule will vary from one firm to another depending on the nature of its business and the attitude of the management toward the assumption of risk. The simplest case would be the firm which has such a powerful aversion to the risks associated with debt that it will under no circumstances engage in debt financing. Such a firm might have a supply-of-funds schedule such as that depicted in Panel A of Figure 10–1. In this case, the amount of retained earnings and depreciation allowances (internal funds) is *OA*. The cost of using internal funds to finance investment is the foregone opportunity cost of investing the funds in the securities of other companies. Assuming that the firm is averse to the risks of playing the stock market, any investment alternative to capital expenditures will be in the debt instruments of other companies (or the government) and the opportunity cost (*OB* in Panel A of Figure 10–1) may be taken as approximately equal to the interest rate the firm would have to pay if it issued its own debt, adjusted for the fact that interest earnings are subject to the corporate income tax. If the interest rate is 6 percent and the corporate income tax rate is 50 percent, the opportunity cost of internal funds will be 3 percent, as shown in Panel A of Figure 10–1.

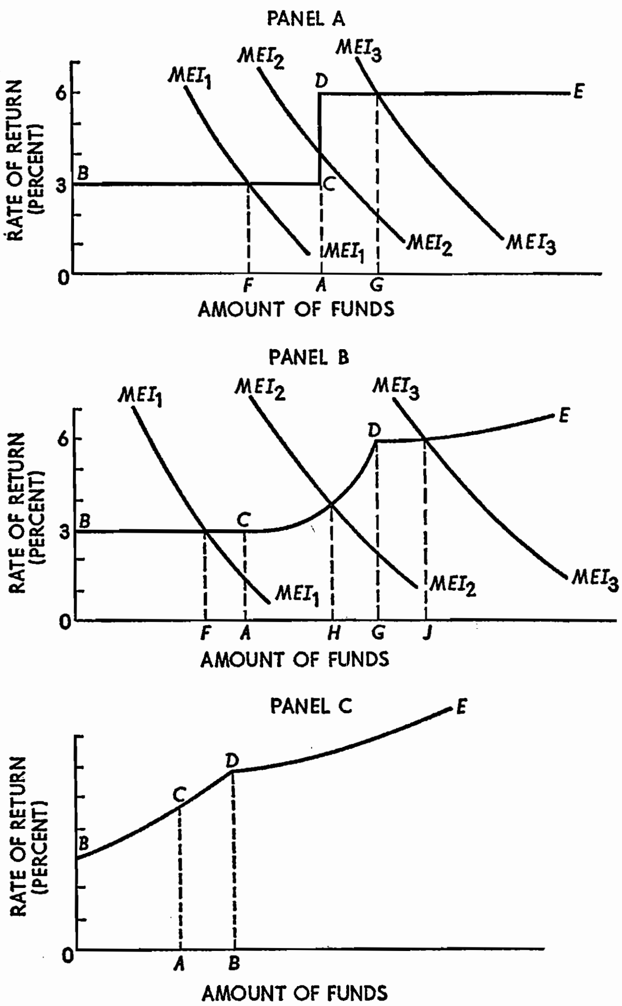
The amount of internal funds (*OA*) is equal to depreciation on the existing capital stock plus retained earnings, which are determined by the application of the firm's dividend policy to its aftertax profits. That is, the amount of internal funds available is assumed to be independent of the opportunities for profitable investment. This is, to some degree, an oversimplification, because the firm might reduce its dividends to make more funds available for investment. This would involve some costs to be balanced against the return from investment, since a reduction in dividends would depress the market value of the firm's stock and might create resentment among the stockholders. In any case, we assume that such adjustments in dividend policy are not made.

If the present stockholders of the corporation are to benefit as a result of the sale of new shares, the proceeds must yield a rate of return greater than the ratio of present average expected earnings to the price of shares. Thus, on the assumption that additional shares can be issued at the current market

---

[3] This discussion of the financing of investment draws heavily on J. S. Duesenberry, *Business Cycles and Economic Growth* (New York: McGraw-Hill Book Co., 1958), chap. 5.

FIGURE 10–1



price, the current earnings-price ratio is a measure of the cost of raising funds through the sale of stock. In Panel A of Figure 10–1, this cost is assumed to be 6 percent, and the line segment *DE* represents the portion of the supply schedule of funds that is derived from the sale of equities.

On the assumption that the firm will not raise any funds by issuing debt, the supply-of-funds schedule is the broken line *BCDE*. If the marginal efficiency of investment schedule is MEI_1, investment will be *OF* and it will be financed entirely by internal funds. If the *MEI* schedule intersects the supply-of-funds schedule anywhere in its vertical segment *DC*, as is the case with MEI_2, investment will be *OA* and will absorb all of the corporation's internal funds. With investment schedule MEI_3, which intersects segments *DE* of the supply-of-funds schedule, investment will be *OG*, of which *OA* will be financed with internal funds and *AG* will be financed by the sale of new equities.

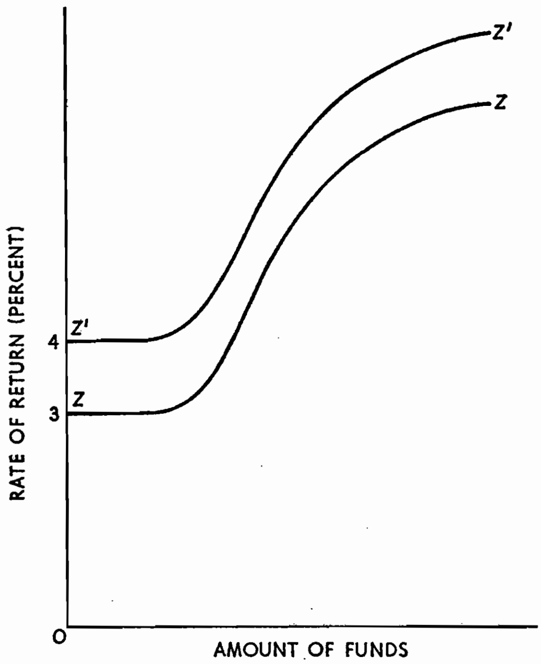
The corporation depicted in Panel B of Figure 10–1 is assumed to have no debt outstanding at the beginning of the period but to be willing to borrow if necessary. The supply-of-funds schedule is *BCDE*. Funds available from internal sources are *BC*. The curved segment *CD* represents borrowing; it is assumed that in addition to the interest that must be paid, the firm imputes additional costs attributable to the risks associated with borrowing that rise as the amount of debt increases. When the explicit interest cost, together with the imputed costs, rises to 6 percent at point *D* (at which point borrowing amounts to *AG*), it becomes worthwhile to finance by selling new stock. In Panel B, allowance is made for the fact, neglected in Panel A, that the sale of additional stock is likely to cause a decline in the price of stock. Thus, the segment *DE* of the supply schedule also slopes upward. In this case, the financing of investment with the three investment demand schedules, MEI_1, MEI_2, and MEI_3 shown in Panel B of Figure 10–1 would be as follows:

MEI Schedule	*Internal Financing*	*Borrowing*	*Sale of Stock*
MEI_1	*OF*	none	none
MEI_2	*OA*	*AH*	none
MEI_3	*OA*	*AG*	*GJ*

In Panel C of Figure 10–1, it is assumed that the firm has some debt outstanding at the beginning of the period. If this is the case, internal funds, which are assumed to be *OA*, could be used to retire a portion of this debt. Since debt is assumed to involve risks to the firm, the use of internal funds to finance capital expenditures involves increasing costs associated with foregone debt retirement. The segment *BC* of the cost of funds schedule, involving the use of internal funds, therefore slopes upward.

The supply-of-funds schedules of individual firms will vary in position and shape, depending on the amount of internal funds available, the attitude of the management toward borrowing and the risks associated therewith, and so on. However, the schedules in general are likely to be similar in appearance to those shown in Figure 10–1. When the schedules of individual firms are combined, the resulting aggregative supply schedule might look like curve ZZ in Figure 10–2. That is, it is likely to be somewhat S-shaped, with an elastic portion at the lower end where internal sources of funds are dominant, an inelastic portion in the middle range as borrowing becomes more important, and a more elastic portion again at the upper end where equity financing predominates. An increase in the prevailing rate of

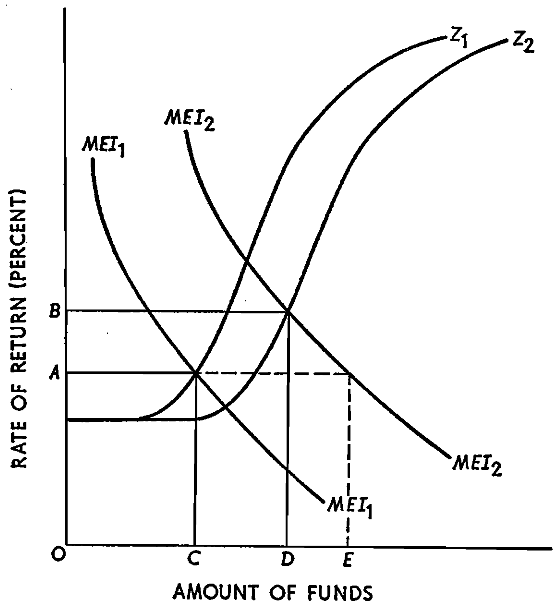
FIGURE 10–2



interest would raise the supply curve vertically, since it would increase the opportunity cost of internal funds and the explicit cost of borrowing and would cause the price of equities to fall to restore an appropriate risk margin between the interest rate on bonds and the earnings/price ratio of equities. Thus, a rise in the interest rate, adjusted for taxes, from 3 to 4 percent might cause the supply schedule to shift from ZZ to $Z'Z'$ in Figure 10–2.

An increase in GNP that raised corporate profits and therefore retained earnings, assuming the interest rate was unchanged, would shift the supply schedule to the right by adding to the elastic portion at the lower end derived from internal funds. That is, an increase in GNP might shift the supply curve from Z_1 to Z_2 in Figure 10–3.

FIGURE 10–3



THE INTERACTION OF INVESTMENT DEMAND AND THE SUPPLY OF FUNDS

As was pointed out earlier, during a rapid expansion in economic activity both investment demand and the supply of internal funds are likely to increase. However, investment demand will generally increase more than the supply of internal funds. Thus, during an expansion, investment demand might increase from MEI_1 to MEI_2 in Figure 10–3 while the supply of internal funds was increasing from Z_1 to Z_2. As a result the marginal cost of funds might be driven up from *OA* to *OB* as firms found it necessary to increase their borrowing. The rising cost of funds would choke off some investment, causing investment spending to rise only from *OC* to *OD* rather than all the way to *OE* as would have occurred had the cost of funds not risen. Curves Z_1 and Z_2 are drawn on the assumption that the interest rate prevailing in the market does not rise; such a rise in the interest rate would be likely to occur during a period of expansion and would exert a further braking effect on investment by causing the curve Z_2 to shift upward.

It is quite clear that the presence of accelerator-type responses introduces an inherent, built-in element of instability in the economy. However, as was shown earlier, the presence of lags in the process of bringing the desired capital stock into line with the actual capital stock serves to damp the fluctuations that occur. It is now apparent that the rising cost of financing will serve to provide a further dampening effect, thereby adding to the stability of the economy. The student should also recall the automatic stabilizers discussed in Chapter 6, which serve to reduce the multiplier, thereby muting the total effect on income and employment of such fluctuations in investment as do occur.

The fact that the four recessions of the postwar period—those of 1949, 1953–54, 1957–58, and 1960–61—have all been quite mild and short-lived suggests that the forces of instability, while present in the U.S. economy, are quite strongly muted. This does not, however, mean that the economy can be relied upon to achieve an adequate performance automatically without the careful guidance of suitable fiscal and monetary policies. As will be shown in Parts III and IV, even if the inherent stability of the economy is sufficient to avoid severe and protracted recessions, the problem of maintaining an appropriate balance between the goals of high employment and price stability in an economy whose productive capacity is continuously increasing certainly requires a vigorous and skillful use of stabilization policies.

EMPIRICAL EVIDENCE ON THE DETERMINANTS OF INVESTMENT

In the last few years, a great deal of empirical work has been undertaken for the purpose of isolating the determinants of investment. The approach employed has generally been to specify models of investment behavior in mathematical form and to fit them to historical time-series data by means of regression techniques. Although the models used differ considerably in detail, nearly all of them fit in a general way within the framework developed above. That is, the typical model incorporates some kind of flexible version of the acceleration principle, together with the interest rate as a measure of the cost of external funds and retained earnings plus depreciation allowances as a measure of funds generated internally.[4] It is true that there has been some dispute concerning the role of internal funds as an independent influence. Some students of investment feel that the firm should in principle be indifferent between funds derived from internal and external sources and that the acceleration principle by itself should provide an adequate explanation of investment.[5] According to this view, the desired stock of capital should depend on the expected level of output and sales and

[4] For a good discussion of the issues and a general summary of results, see Edwin Kuh, "Theory and Institutions in the Study of Investment Behavior," *American Economic Review,* Vol. LII, May 1963, pp. 260–68. An interesting discussion focused around the problem of forecasting is to be found in Avram Kisselgoff, "Forecasting Investment," in *The Economic Outlook for 1968,* Papers presented to the 15th Annual Conference on the Economic Outlook at the University of Michigan (Ann Arbor: Department of Economics, University of Michigan, 1968), pp. 145–53.

[5] See Robert Eisner, "Investment: Fact and Fancy," *American Economic Review,* Vol. LIII, May 1963, pp. 237–46.

on the interest rate, with investment determined by a gradual process of adjusting the actual stock to bring it into line with the desired stock along the lines of the stock adjustment mechanism described in Chapter 9. In general, however, the models that have been most successful in explaining investment behavior have generally included the flow of internal funds as a separate variable, and in some instances they have also incorporated other aspects of the financial position of the firm such as the amount of debt outstanding.[6]

One of the main problems in the estimation of equations to explain investment is the proper specification of lags. In the stock adjustment model discussed earlier, the adjustment of the actual capital stock to bring it to equality with the desired capital stock was spread out over time, with a certain fraction of the gap being filled by investment in each period. No explanation was given to account for the existence of the lags in the adjustment process. One possible explanation is as follows.[7] Suppose that when a gap between the actual and desired capital stock appears, businessmen place sufficient new orders for capital goods to fill the gap. That is, their orders are sufficient to make the backlog of unfilled orders equal to the gap. However, many steps lie between the placing of an order for a capital good and the final expenditure on the good which is recorded as investment. Engineering plans must be drawn up, orders placed, and the structures or equipment must be produced. Suppose that of any given backlog of unfinished work, a certain part is finished and paid for in the first period, a further portion is completed and paid for the next period, and so on. If a constant and specified fraction, s, of the remaining order backlog is completed and paid for in each period, the time pattern will be of the kind indicated in the stock adjustment model developed earlier in this chapter.[8]

Other lag patterns are possible, however. For example, any backlog might be worked off in n periods with a fraction equal to $1/n$ being completed and paid for in each period.[9] Or completions might be small at

---

[6] See, for example, Frank deLeeuw, "The Demand for Capital Goods by Manufacturers: A Study of Quarterly Time Series," *Econometrica,* Vol. XXX, July 1962, pp. 407–23; W. H. L. Anderson, *Corporate Finance and Fixed Investment* (Boston: Graduate School of Business Administration, Harvard University, 1964), and "Business Fixed Investment: A Marriage of Fact and Fancy," in Robert Ferber (ed.), *Determinants of Investment Behavior* (New York: National Bureau of Economic Research, 1967), pp. 413–25.

[7] This is the approach to the lags of adjustment that is used by deLeeuw, *op. cit.*

[8] For example, suppose to begin with that the actual and desired capital stock are equal and that in the next period a gap of 100 is opened up and 100 of orders are placed to fill the gap. If s is 20 percent, 20 of capital goods will be completed after one period. If the desired stock of capital remains unchanged so that no more orders are placed, 20 percent of the remaining 80 of backlog, or 16, will be completed and paid for after two periods, 20 percent of the remaining 64, or 12.8, the next period, and so on. Thus, the sequence of investment expenditures in successive periods will be 20, 16, 12.8, etc.

[9] In the example of footnote 8, the total backlog might be worked off in five periods, with one fifth being completed in each period. Then investment would be 20 per period for five periods.

first, then build up to a peak and taper off.[10] The lag pattern depends primarily on the production processes involved in manufacturing capital goods and the composition of the order backlog—for example, buildings may take longer to produce than machinery. Usually the lags are estimated as part of the statistical process of fitting the investment equation to the empirical data—that is, the lag pattern selected is the one that provides the best fit to historical data.

A variety of models have been tested, and several of them are sufficiently successful in explaining past investment behavior that it is scarcely justifiable to select one as clearly superior to the rest. Nevertheless, some progress in explaining investment seems to have been made, and some generalizations are possible. In most instances, the more successful models involve an application of some variant of the acceleration principle with the actual capital stock being adjusted gradually to bring it into line with the desired capital stock. Most of the studies also show that the interest rate affects investment, often through its influence on the desired capital stock. The flow of internal funds generally appears as an additional explanatory variable, indicating that businessmen are not indifferent as between internal and external sources of funds. Most studies show that the lags are quite long, that is, it takes many quarters to bring about a full adjustment of the stock of capital. Moreover, the preponderant evidence seems to be that the lag pattern is one in which the response is small at first, builds up to a peak, and then gradually tapers off.

One particular issue that is of considerable importance is the degree of sensitivity of investment to changes in the rate of interest. Monetary policy is generally thought to affect the economy primarily by producing changes in the interest rate, and its effectiveness as an economic regulator may therefore depend critically upon the sensitivity of investment to changes in the interest rate.

To begin with, it may be noted that there are a priori reasons for doubting that investment decisions would be likely to be strongly affected by changes in borrowing costs. One is that interest is deductible as an expense in calculating income for tax purposes. With a corporate income tax rate of 50 percent, an increase in the interest rate by one percentage point, from 5 to 6 percent, would mean an increase of only half a percentage point, from 2.5 to 3 percent, in the *effective* interest rate after adjustment for tax deductibility. Thus, the deductibility of interest under the income tax probably serves to blunt rather substantially the effects of interest rate changes on investment. In addition, and more fundamental, there is the inherent element of risk involved in investment. As was brought out in Table 8–1, investment is a forward-looking activity whose prospective profitability is

[10] As a simple illustration, the backlog of 100 specified in footnote 8 might be worked off according to a so-called "triangular" or "inverted V" pattern over eight periods, with investment for the successive periods being 5, 10, 15, 20, 20, 15, 10, 5.

dependent on the investor's expectations concerning sales, prices, wages, and so on, over the life of the investment. Relatively small errors in forecasting these elements can have large effects on the investor's profits from the investment. For this reason, the investor may inject a rather large allowance for risk into his calculations. Thus, he may not be willing to undertake an investment unless he expects to earn a return of, say, 15 percent on his money after taxes. If this is the case, an increase of 1 percent in the interest rate—a rather large increase by historical standards—which, assuming a 50 percent tax rate, would raise his effective borrowing cost by only one half of 1 percent, might well be an insignificant factor in his calculations. Surveys of businessmen's views regarding the effects of interest rates on investment decisions have generally seemed to support the hypothesis that interest rates are not very important.[11]

The common approach in econometric studies of investment at the present time is to assume, in effect, that the desired stock of capital depends upon the interest rate as well as upon income. If this is the case, a reduction in the interest rate increases the desired stock of capital, thereby setting off a sequence of increments to investment designed to raise the stock of capital gradually toward the new desired level. With the kinds of lags that are present in most of the models, the effects of a change in the interest rate on investment are small at first, building up to a peak, and then declining. Although the exact nature of the results varies from one study to another, most of the recent work suggests that the interest rate has a fairly sizable effect on investment but that several quarters must elapse before a very significant portion of the effect is felt.[12]

In addition to their effects on business investment in plant and equipment, interest rates and credit conditions may have important effects on other components of final demand that are commonly financed by the use of credit, including residential construction; construction of schools, highways, and other public projects by state and local governments; the demand for consumer durable goods; and inventory investment. The responses of these types of spending to monetary conditions are often greatly affected by the specific financial arrangements that influence the flows of credit to these sectors. Detailed analysis of these effects would require an examination of the financial sector of the economy that is well beyond the scope of this book.[13]

---

[11] The results of a recent survey of this kind are reported in Jean Crockett, Irwin Friend, and Henry Shavell, "The Impact of Monetary Stringency on Business Investment," *Survey of Current Business,* August 1967, pp. 10–27.

[12] On this, see Albert Audo, E. C. Brown, R. M. Solow, and John Kareken, "Lags in Fiscal and Monetary Policy," in *Stabilization Policies,* Research Study No. 1, Commission on Money and Credit (Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1963), pp. 1–7 and 14–96.

[13] The preliminary results derived from a rather complex aggregative model of the economy especially designed to uncover the effects of interest rates and monetary condi-

THE ROLE OF INVESTMENT IN THE DETERMINATION OF EQUILIBRIUM INCOME

As the above discussion makes clear, investment is, in essence, a dynamic phenomenon. There is strong reason to believe, based on the theoretical reasoning supported by empirical research, that a stock adjustment mechanism constitutes a primary force governing investment spending. When the economy is in equilibrium, with income constant and the stock of capital optimally adjusted to that level of income, the stock adjustment mechanism will become inoperative as a generator of net investment. This is the case in the "initial equilibrium" and "new equilibrium" columns of Table 9–3, the only investment in these columns being replacement investment amounting to 2 percent of the equilibrium capital stock.

Even in equilibrium, however, some net investment would be generated by the autonomous forces of population growth and technological progress discussed in Chapter 9. At this point, it will be useful to explore the role of investment in the determination of the equilibrium level of income. Since this analysis will be static, we will abstract from the dynamics of the investment accelerator and assume that there exists a static marginal efficiency of investment schedule such as that depicted in Figure 8–2, which relates equilibrium investment to the rate of interest, the position of this schedule being determined by autonomous forces affecting investment. Figure 10–4 shows how the marginal efficiency of investment schedule interacts with the consumption function to establish a relation between equilibrium income and the interest rate.

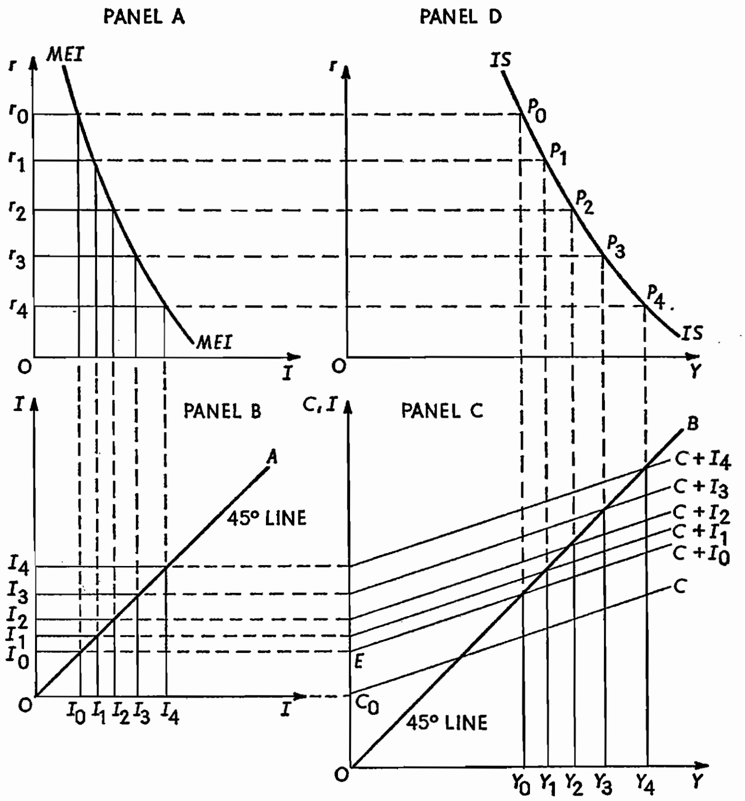
Figure 10–4 is to be studied in counterclockwise fashion, beginning with Panel A in the upper left-hand corner and proceeding successively to Panels B, C, and D. Panel A contains the marginal efficiency of investment (MEI) schedule; if the interest rate is given, this schedule determines the level of investment (I). Panel B shows investment on both axes with a 45° line (line OA) drawn between them, and merely serves to convert investment from a horizontal distance as in Panel A to a vertical distance as required in Panel C. In Panel C, consumption expenditures (C) and investment expenditures are shown on the vertical axis, and income (Y) is shown on the horizontal axis. The 45° line (line OB) in Panel C is thus the locus of points

tions on various components of aggregate demand are reported in Frank deLeeuw and Edward Gramlich, "The Federal Reserve–M.I.T. Econometric Model," in *The Economic Outlook for 1968,* Papers presented to the 15th Annual Conference on the Economic Outlook at the University of Michigan (Ann Arbor: Department of Economics, University of Michigan, 1968), pp. 51–109, reprinted in *Federal Reserve Bulletin,* January 1968, pp. 11–40. Other reports on the progress of this project are given in R. H. Rasche and H. T. Shapiro, "The F.R.B.–M.I.T. Econometric Model: Its Special Features," *American Economic Review,* Vol. LVIII, May 1968, pp. 123–49; and Albert Ando and Franco Modigliani, "Econometric Analysis of Stabilization Policies," *American Economic Review,* Vol. LIX, May 1969, pp. 296–314.

at which expenditures (measured vertically) and income (measured horizontally) are equal. The line C_0C in Panel C is the consumption function, of the short-run type depicted in Figure 5–1. The distance OC_0 indicates the level of the consumption function, and Panel C is located in such a way that

FIGURE 10–4

Integration of Marginal Efficiency of Investment Schedule and Multiplier Analysis



the horizontal axis of Panel B coincides with point C_0 on the vertical axis of Panel C. Panel D will be explained below.

If the interest rate is given at r_0 as shown in Panel A, investment will be I_0 as determined by the *MEI* schedule at that interest rate. Tracing through Panel B, investment I_0 is converted from a horizontal distance into a vertical distance, which is carried over to the vertical axis of Panel C at point *E*. From point *E* a line is drawn parallel to the consumption function. This line, labeled $C + I_0$, indicates the level of total expenditures (consumption plus investment) associated with each level of income when investment is I_0. The equilibrium level of income associated with investment I_0 is designated Y_0 and is determined by the intersection of the line $C + I_0$ and the 45° guideline (line OB) in Panel C. Thus, if the *MEI* schedule and the consumption function are as indicated in Panels A and C respectively, an income of Y_0 will be associated with the interest rate r_0.

By the same reasoning as that used above, an income level of Y_1 will be associated with interest rate r_1, Y_2 with r_2, Y_3 with r_3, and so on. The results are summarized in Panel D, which is constructed by tracing horizontal lines from Panel A at each interest rate and tracing vertical lines from Panel C for the corresponding income levels. The intersections of the horizontal lines at each interest rate with the vertical lines at each corresponding income level determine a series of points (P_0, P_1, P_2, P_3, and P_4). The line *IS* in Panel D passes through all of these points and thus sums up the relation between interest rate and income level.

This line, passing through the equilibrium points in Panel D, is designated the "*IS* curve" because it contains all combinations of income and interest rate for which planned investment (I) is equal to planned saving (S). Or, to put it another way, the *IS* curve includes all combinations of Y and r for which aggregate demand for goods and services equals aggregate supply.

It should be reemphasized that the analysis presented in Figure 10–4 makes no allowance for the operation of the stock adjustment mechanism, which was given the central role in explaining fluctuations in the position of the marginal efficiency of investment schedule in Chapter 9. In the present analysis, the position of the *MEI* schedule is taken as given, determined by the autonomous forces of population growth and technological change, which are assumed to be the underlying determinants of the equilibrium rate of investment. It is not possible to incorporate the stock adjustment mechanism which is inherently dynamic within the present static framework. That is, the analysis of Figure 10–4 is entirely *static* and merely serves to describe alternative *equilibrium* combinations of income and interest rate, without specifying the way in which the economy moves from one equilibrium position to another. And since the level of income is constant in equilibrium, there can be no investment induced by accelerator or stock adjustment mechanisms in the present analysis.

In the model depicted in Figure 10–4, the level of income can change for essentially three reasons. First, the interest rate, which in the present context is an exogenous variable (i.e., determined by forces not included in the model and therefore subject to independent variation), may change, thus causing movement along the *IS* curve in Panel D. Second, the equilibrium position of the *MEI* schedule in Panel A may shift due to technological change or for other reasons. This will cause the *IS* curve in Panel D to shift and if the interest rate remains constant, income must change. Third, the consumption function, as shown in Panel C, may shift for various reasons discussed in Chapter 5, and, again, if the interest rate remains unchanged, income must change. The student may verify for himself by studying Figure 10–4 that a shift to the right of the *MEI* schedule in Panel A or an upward shift on the consumption function (the line C_0C) in Panel C will cause the *IS* curve in Panel D to shift to the right. Conversely, a shift to the left of the *MEI* schedule or a decline of the consumption function will cause the *IS* curve to shift to the left.[14]

[14] It may be noted that a rise in the consumption function (the line C_0C in Panel C) will be depicted in Figure 10–4 by a lowering of the *horizontal axis* of Panel C, since Panel C must be placed so that the point C_0 lines up with the horizontal axis of Panel B. Similarly, a fall in the consumption function will be depicted by a raising of the horizontal axis of Panel C.

Chapter 11 THE DEMAND FOR MONEY

Up to this point we have paid little attention to money and the financial sector of the economy. It is now time to remedy this defect. In this chapter we shall consider the demand for money and explore the role of monetary forces in determining the interest rate.

RELATION BETWEEN BOND PRICES AND INTEREST RATES

As a preliminary to our analysis, it is necessary to explain the relationship that exists between interest rates and the prices of existing marketable bonds.

A *bond* is a contractual debt obligation which may be issued by a business corporation or a governmental unit and which promises to pay the holder certain periodic interest payments and to pay the *principal value* at maturity.[1] The interest payments are ordinarily calculated by applying the interest rate specified on the bond, called the *coupon rate,* to the principal value. Thus, a $1,000, 3 percent, 20-year bond, with interest payable annually, has a principal value of $1,000 and a coupon rate of 3 percent. Such an instrument, at the time of its original issuance, promises to pay $30 interest (3 percent of $1,000) each year for 20 years and also to pay back the $1,000 principal at the end of 20 years.

It is very important to understand that the coupon rate on a bond is fixed by the original bond contract and does not change during the life of the bond. However, the interest rate that borrowers must pay when they raise funds through the issuance of additional bonds changes continually in response to changes in the demand for and supply of such bonds. Thus, the coupon rate that a particular corporation must place on 10-year bonds to

[1] There are many different kinds of bonds, which may be classified, for example, according to the kind of protection afforded the bondholder. These matters are not relevant to our discussion, and for further information concerning them the reader is referred to any good book on corporation finance or on investments, such as L. V. Plum and J. H. Humphrey, Jr., *Investment Analysis and Management* (Homewood, Ill.: Richard D. Irwin, Inc., 1953).

induce investors to buy them at their face (or principal) value may be, say, 3 percent at one time and 5 percent at another, depending upon the state of business, the monetary policies being followed by the Federal Reserve authorities, and a variety of other factors. At the same time, the return on existing marketable bonds of the same quality (i.e., judged by investors generally to involve the same risk of default on interest and principal) issued at an earlier time and now having 10 years to run to maturity will tend to change in such a way as to remain in line with the return to be had from buying new bonds. Since the coupon rate on existing bonds is fixed by contract, the adjustment takes the form of a change in the price of such bonds.

To illustrate, suppose we take an existing bond having a principal value of \$1,000, a coupon rate of 5 percent, interest payable annually, and one year to run to maturity; and suppose the current interest rate on similar bonds is 4 percent. The existing bond is essentially a contract to pay \$1,000 (principal) plus \$50 (interest) or \$1,050 in total at the end of one year. We can calculate the price (V) that a knowing investor would be willing to pay for the existing bond in the following way:

$$V(1.04) = \$1{,}050$$

or

$$V = \frac{\$1{,}050}{1.04} = \$1{,}009.62$$

That is, the price can be determined by finding out how much would have to be invested at the prevailing interest rate in order to build up to \$1,050 at the end of one year.

This is essentially the same discounting procedure that was discussed in Chapter 8 in analyzing the concept of marginal efficiency of investment.[2] The general formula for the price of a bond is

$$V = \frac{cP}{1+r} + \frac{cP}{(1+r)^2} + \cdots + \frac{cP}{(1+r)^n} + \frac{P}{(1+r)^n} \qquad (1)$$

where V is the market price of the bond, c is the coupon rate (expressed as a decimal) per interest payment period (ordinarily a half year or a year) on the bond, P is the principal value of the bond, r is the current market rate of interest (expressed as a decimal) per interest payment period on bonds of the same maturity as the one being considered, and n is the number of interest payment periods in the remaining life of the bond. In the simple example discussed above, the interest payment period is one year (i.e., interest is paid annually), $c = 0.05$, $P = \$1{,}000$, $r = 0.04$, and $n = 1$.

All of the terms in Equation 1 before the last term represent the present

[2] See Chapter 8, pp. 164–67.

(discounted) value of future interest payments—i.e., the first term $(cP/(1 + r))$ is the present value of the first interest payment to be received at the end of one year, the second term $(cP/(1 + r)^2)$ is the present value of the second interest payment to be received at the end of two years, and so on. By the same reasoning that was used in Chapter 8,[3] the sum (K_n) of the present values of all the interest payments (all terms in Equation 1 above except the last term) is given by

$$K_n = \frac{cP}{r}\left[1 - \frac{1}{(1 + r)^n}\right]$$

Adding in the last term of Equation 1, the formula for V may be rewritten as

$$V = \frac{cP}{r}\left[1 - \frac{1}{(1 + r)^n}\right] + \frac{P}{(1 + r)^n} \qquad (2)$$

If the terms of a bond—coupon rate (c), principal value (P), and time to maturity (n)—are known, this formula can be used to find (1) the interest rate (r) or "yield" that will be earned if the bond is purchased at a specified price (V), or (2) the price that must be paid for the bond if it is to provide the investor with a specified interest return or "yield." The formula is mathematically cumbersome, but it need not actually be solved by the user since so-called "bond tables" are available in which the values of V are tabulated for various combinations of c, r, and n for a bond having a face value of one dollar ($P = 1$). By using these tables, it is possible quite easily to find the yield corresponding to a given price or the price corresponding to a given yield for any specified bond.[4]

By studying Equations 1 and 2—which, of course, are merely two different ways of stating the same relationship—it is possible to draw several conclusions concerning the relation between the market rate of interest (r) and the price of an existing bond (V).

1. The prices of existing bonds fall when the interest rate rises and rise when the interest rate falls.[5] This is apparent from the fact that a rise in the

[3] See Chapter 8, pp. 166–67.

[4] Since bond tables give prices corresponding only to discrete maturity intervals (months, for example) and yield intervals (perhaps one 10th of 1 percent), it is often necessary to interpolate where the maturity of the bond falls between two of the dates or its yield between two of the interest rates that are specifically shown in the table. In addition, there are complications in computing yields or prices arising out of special provisions in bond contracts, such as "call features" which permit the issuer of the bonds to retire them on certain specified dates prior to final maturity if he chooses. For a brief but informative explanation of such matters—which are of little importance for our purposes—see Plum and Humphrey, *op. cit.*, Appendix 2, pp. 630–40. The standard bond tables include *Executive Bond Values Table* (Boston: Financial Publishing Co., 1947), and *Comprehensive Bond Values Table* (Boston: Financial Publishing Co., 1958).

[5] The prices at which newly issued bonds sell also fluctuate with current interest rates. This is because the coupon rate that is placed on a bond by its issuer seldom corresponds

interest rate (r) increases the denominators in each of the terms of Equation 1, thus reducing the value of each term and lowering the price (V) of the bond. A fall in the interest rate has the opposite effect.

2. If the current interest rate (r) is equal to the coupon rate (c), the price of the bond (V) is equal to its principal value (P), or the bond is said to sell at "par," regardless of the maturity of the bond. In this case, $c = r$, and on substituting r for c, Equation 2 becomes

$$V = \frac{rP}{r} 1 - \frac{1}{(1+r)^n} + \frac{P}{(1+r)^n}$$

or

$$V = P - \frac{P}{(1+r)^n} + \frac{P}{(1+r)^n}$$

or

$$V = P$$

3. It follows from the two propositions stated above that if the current interest rate is below the coupon rate ($r < c$), the bond will sell above par ($V > P$); and if the current interest rate is above the coupon rate ($r > c$), the bond will sell below par ($V < P$).

4. The longer the time to maturity of the bond (i.e., the greater the value of n), the greater will be the change in its price resulting from a given change in the interest rate. This is apparent from Equation 1, since the more distant a payment is, the more its present value is affected by a change in the rate of interest—that is, a change in r changes $(1 + r)^2$ more than $1 + r$, $(1 + r)^3$ more than $(1 + r)^2$, and so on. Since a long-term bond has terms containing larger powers of $1 + r$, its price is affected more by a change in r than is the price of a short-term bond. As a result of this, the prices of long-term bonds are more unstable than are the prices of shorter term bonds.

5. A *consol* is a special type of bond which continues to pay interest at the contractual rate indefinitely but makes no provision for repayment of the principal. Consols are not very common, at least in this country, but they are frequently used for illustrative purposes, because their behavior is rather similar to that of very long-term bonds in general and because the relation between the interest rate and the price of the bond reduces to a very simple formula in the case of consols. For a consol, the maturity of the bond (n) becomes infinite. The expression $(1 + r)^n$ becomes infinitely large as n ap-

---

exactly with the yield at which investors are willing to buy the bond, so that the bond will ordinarily sell at a premium or discount (i.e., a price above or below par) sufficient to adjust its yield to the requirements of the market. A notable exception is United States government bonds, which are usually issued at a fixed price (commonly par).

proaches infinity, so that the terms $1/(1+r)^n$ and $P/(1+r)^n$ in Equation 2 reduce to zero. Thus, the expression becomes simply

$$V = \frac{cP}{r}$$

for a consol.[6] That is, the price of a consol changes in inverse proportion to the rate of interest.

All five of these propositions are illustrated in Table 11–1, which shows the prices of a $1,000 bond, with a coupon rate of 4 percent, corresponding to various market rates of interest and various assumptions concerning the

TABLE 11–1

Prices of $1,000 Bonds of Various Maturities, Bearing Coupon Rates of 4 Percent, Corresponding to Various Market Rates of Interest*

Current Interest Rate (Percent)	*Years to Maturity* 1	5	10	25	40	Consol ∞
5.0	$ 990.40	$ 956.20	$ 922.10	$ 858.20	$ 827.70	$ 800.00
4.5	995.20	977.80	960.10	925.40	907.60	888.90
4.0	1,000.00	1,000.00	1,000.00	1,000.00	1,000.00	1,000.00
3.5	1,004.90	1,022.80	1,041.90	1,082.90	1,107.20	1,142.90
3.0	1,009.80	1,046.10	1,085.80	1,175.00	1,232.00	1,333.30

*Bonds are assumed to pay interest semiannually.

maturity of the bond. When the market rate of interest is 4 percent—i.e., equal to the coupon rate—the bond sells at par ($1,000) no matter what its maturity. As the interest rate rises above 4 percent, the price of the bond falls; and as the rate falls below 4 percent, the price rises. The longer the maturity of the bond, the greater the fluctuations in its price; for example, a rise in the interest rate from 4 to 4.5 percent causes the price of a bond of five years' maturity to fall from $1,000 to $977.80, while the same rise in the interest rate causes a bond of 25 years' maturity to fall from $1,000 to $925.40. The last column of the table shows the behavior of the price of a consol, which, of course, shows the greatest price fluctuations of all.

THE DEMAND FOR MONEY

Interest is best viewed simply as the price paid for the loan of a specified sum of money for a specified period of time. Thus, the determination of the

[6] This is, of course, an application of the so-called "capitalization formula," which was referred to in Chapter 8, p. 167.

rate of interest may be analyzed in a relatively straightforward fashion by simply studying the interaction between the demand for money and the supply of money. Let us now consider the demand for money, approaching the subject by considering separately the various reasons why economic units choose to hold money balances.

The Motives for Holding Money

It has become fairly standard practice to distinguish three major motives for holding money balances: the transactions motive, the precautionary motive, and the speculative motive.[7]

The Transactions Motive. A considerable amount of money is held simply to bridge gaps between the receipt of money payments and the making of money expenditures. That is, an office worker may receive his salary monthly while his payments for groceries, rent, insurance, clothing, etc., are spread throughout the month in a manner determined by his family's habits, contracts previously entered into, and to some extent random or accidental factors. Given the timing of his receipts and expenditures, the worker will hold a certain average cash balance for transactions purposes in relation to his income and expenditures.[8]

Some of the factors that determine the holdings of transactions balances in relation to income may be brought out by considering some simple examples.[9] Suppose for instance, that a worker receives an income of $3,600 per year, that he spends all of his income, and that he spends at a steady rate of $10 per day. If he receives his entire annual income at the beginning of

[7] This way of classifying the motives for holding money was first used by J. M. Keynes in *The General Theory of Employment Interest, and Money* (New York: Harcourt Brace & Co., 1936).

[8] Money balances are held for the purpose of financing transactions in securities as well as in goods and services. However, our analysis of the transactions motive is confined to goods and services transactions. The cash balances used to handle securities transactions (the so-called "financial circulation") are largely confined to professional dealers and traders, and methods of economizing cash balances have been perfected to such a high degree in this connection that the billions of dollars of financial transactions that are carried out each year require only very small amounts of cash balances. For example, in February 1959, according to the New York Clearing House Association, dealers in United States government securities and other investment dealers and brokers wrote checks totaling $28 billion while carrying average bank balances of $406 million. Thus, the rate of turnover of their deposits was 858 times per year, compared with 32 per year for other depositors. The turnover of deposits for dealers in United States government obligations was at the fantastic rate of 11,264 times per year! Data are from *Member Bank Reserve Requirements,* Hearings, Committee on Banking and Currency, U.S. Senate, 86th Congress, 1st Session, on S860 and S1120, March 23–24, 1959, pp. 172 and 89.

[9] For a much more sophisticated analysis of these matters, see J. W. Angell, "The Components of the Circular Velocity of Money," *Quarterly Journal of Economics,* Vol. LI, February 1937, pp. 224–73; and H. S. Ellis, "Some Fundamentals in the Theory of Velocity," *Quarterly Journal of Economics,* Vol. LII, May 1938, pp. 431–72, reprinted in F. A. Lutz and L. W. Mints (eds.) *Readings in Monetary Theory* (Philadelphia: Blakiston Co., 1951), pp. 89–128.

the year, he will have $3,600 in his cash balance on January 1 and nothing at the end of the year. Since his balance declines steadily day by day, clearly his *average* balance will be $1,800, which is equal to one half of his annual income. Now if this person maintains the same pattern of payments, $10 per day, but is paid $300 at the beginning of each month, his average balance will fall steadily from $300 at the beginning of each month to zero at the end of each month, and his average cash balance will be $150 or one 24th of his annual income of $3,600. That is, when he is paid monthly, his average balance is one 12th as great as it is when he is paid annually. Thus, other things remaining constant, the average balance declines as the *frequency of receipts* increases.

Suppose now that the worker receives a salary of $300 at the beginning of each month but that instead of spending steadily so much each day, he spends $100 on the 10th of the month, another $100 on the 20th, and the final $100 on the 30th. Then his average cash balance can be calculated as follows:

$300 per day for 10 days	$3,000
$200 per day for 10 days	2,000
$100 per day for 10 days	1,000
Total	$6,000
Average balance ($6,000 ÷ 30 days)	$200

In this case, the worker's average cash balance is one 18th of his annual income of $3,600 compared with a balance of $150 or one 24th of annual income in the case in which he spends his income steadily day by day. Thus, the average cash balance declines as the *frequency of expenditures increases.*

A third factor on which the average cash balance depends is the extent to which receipts and expenditures coincide. This can be seen by assuming that the worker receives $300 per month and spends $100 every 10 days, as in the previous example, but making a slight change in the pattern. Suppose the worker spends $100 as soon as he receives his income, another $100 on the 10th of the month, and the final $100 on the 20th—that is, each of the $100 expenditures is moved forward by 10 days as compared with the previous example. In this case, his average cash balance can be computed as follows:

$200 per day for 10 days	$2,000
$100 per day for 10 days	1,000
$ 0 per day for 10 days	0
Total	$3,000
Average balance ($3,000 ÷ 30 days)	$100

The average balance is one 36th of annual income, as compared with one 18th in the previous example. Thus, the average balance is smaller, the greater is the *coincidence between receipts and expenditures.*

Although the above analysis is somewhat oversimplified, it does suggest that we may view the ratio of transactions balances to expenditures on goods and services for a spending unit as depending upon that unit's frequency of receipts, frequency of expenditures, and degree of coincidence between receipts and expenditures. Looking now at the economy as a whole, we may define k_t as the ratio of cash balances held to finance expenditures on goods and services to the total amount of such expenditures, M_t as total cash balances held for transaction purposes, and X as total expenditures on goods and services. Then by definition,

$$k_t = \frac{M_t}{X} \tag{3}$$

We can also define similar concepts for each spending unit in the economy. Thus, $k_{t,1}$, $M_{t,1}$, and X_1 may stand for the ratio of cash balances to expenditures, the amount of cash balances, and the volume of expenditures, respectively, for the first spending unit in the economy; $k_{t,2}$, $M_{t,2}$, and X_2 may stand for similar concepts for the second spending unit, and so on. Then, if there are n spending units in the economy, we have

$$M_t = M_{t,1} + M_{t,2} + \cdots + M_{t,n}$$

and

$$X = X_1 + X_2 + \cdots + X_n$$

Since the k_t's of the individual spending units are defined as $k_{t,1} = M_{t,1}/X_1$, $k_{t,2} = M_{t,2}/X_2$, and so on, we have

$$M_t = k_{t,1}X_1 + k_{t,2}X_2 + \cdots + k_{t,n}X_n$$

Substituting this into Equation 3 above, we obtain

$$k_t = \frac{k_{t,1}X_1 + k_{t,2}X_2 + \cdots + k_{t,n}X_n}{X}$$

or

$$k_t = k_{t,1}\frac{X_1}{X} + k_{t,2}\frac{X_2}{X} + \cdots + k_{t,n}\frac{X_n}{X} \tag{4}$$

Let us now define k_y as the ratio of transactions balances to gross national product (Y). That is,

$$k_y = \frac{M_t}{Y} \tag{5}$$

As was shown in the discussion of national income accounting in Chapter 2, the gross national product (Y) is substantially smaller than total expenditures on goods and services by all spending units (X), because the latter total includes a large amount of intermediate transactions which are canceled out in calculating the GNP in order to avoid double-counting. Thus, we can rewrite Equation 5 as:

$$k_y = \frac{M_t}{X} \cdot \frac{X}{Y}$$

or

$$k_y = k_t \frac{X}{Y}$$

Finally, substituting for k_t the value given by Equation 4, we can write

$$k_y = \left[k_{t,1} \frac{X_1}{X} + k_{t,2} \frac{X_2}{X} + \cdots + k_{t,n} \frac{X_n}{X} \right] \frac{X}{Y} \tag{6}$$

From this last expression, it is apparent that k_y, the ratio of transactions balances to GNP, depends upon (1) the ratios of cash balances to expenditures for the individual spending units (the individual k_t's); (2) the shares of total expenditures carried out by individual spending units (the ratios X_1/X, X_2/X, etc.), which serve as weights to be used in averaging the individual k_t's; and (3) the ratio of total expenditures on goods and services to the GNP (X/Y). Although there is little if any evidence concerning the values of the k_t's of individual spending units, it is customarily assumed—and with a certain amount of plausibility—that the payment patterns (frequency of receipts, frequency of expenditures, and degree of coincidence of receipts and expenditures) on which the k_t's depend are not likely to change significantly in the short run. While the distribution of expenditures among spending units, which serve as the weights of the k_t's in the expression for k_y, undoubtedly change as GNP changes, there is no evidence that they change in a systematic way which would make k_y rise or fall as GNP changes. Similarly, there is no evidence that the ratio of total expenditures to GNP changes systematically as GNP changes in the short run.[10] For these

[10] In the longer run, there are likely to be changes in k_y, but it is not entirely clear what the direction of these changes will be. Systematic changes in payment patterns may alter the k_t's of individual spending units so as to increase or decrease k_y. For example, if there is a systematic shift from payment of wages and salaries weekly to semimonthly or monthly payments, this may raise the k_t's of many people and thus raise k_y. Vertical integration of industry—that is, the combining in one firm of operations formerly conducted by separate firms, such as the taking over of iron ore mining operations by firms producing steel—will reduce intermediate expenditures relative to GNP, thus reducing X/Y in Equation 6 and probably causing k_y to decline. The commonsense interpretation of this is that vertical integration means that some transactions formerly occurring between firms and requiring money payments become transactions between units of the same firm which can be settled by accounting entries. A reduction in money payments is likely to cause a reduction in the transaction demand for money. Such a result is not absolutely certain, however, because vertical integration also changes some of the k_t's and their weights, with results that are not, in general, easily determined.

reasons, it has been rather commonly assumed that k_y is a stable magnitude in the short run and that it is approximately correct to treat it as a constant. That is, it has frequently been assumed that the demand for transactions balances changes in proportion to income.[11]

Recently, however, it has come to be recognized that the demand for cash balances for transactions purposes may depend upon the interest rate as well as the level of income.[12] A spending unit which receives money payments all or part of which will be needed at a later date for the payment of obligations need not necessarily hold the receipts in cash during the period involved. Instead, the spending unit may invest the funds in interest-bearing securities and then disinvest (i.e., sell the securities) before the payment comes due. The advantage of this procedure, of course, is that interest is earned on the securities during the period of investment. There are, however, certain costs that must be set against the interest returns, the most important of which are the fees and commissions that must be paid to dealers in securities at the time the securities are bought and sold.[13] If spending units systematically balance returns against costs in order to minimize the losses attributable to the holding of sterile cash balances, the tendency to invest and disinvest in securities will increase and the average cash balance held at a given level of expenditures will fall as the interest returns obtainable rise relative to the costs of engaging in security transactions.[14] As a result of

---

[11] This assumption is suggested as plausible, for example, by Keynes, *op. cit.*, p. 201.

[12] This possibility was pointed out in A. H. Hansen, *Monetary Theory and Fiscal Policy* (New York: McGraw-Hill Book Co., 1949), pp. 66–68.

[13] Unless there are claims available whose prices are fixed or at least quite stable, the risk of loss from having to liquidate claims at times when prices are unfavorable in order to meet transactions needs will tend to deter investment of transactions balances in interest-bearing debt instruments. However, very short-term debt, such as U.S. Treasury bills maturing in 30 days or less, fluctuate very little in price as interest rates change, for reasons discussed above in connection with Table 10–1, and are therefore quite safe investments for transaction balances.

[14] A simple example will serve to illustrate the possible effect of the interest rate on the demand for transactions balances. Suppose a spending unit receives a payment of $2,000 which it expects to need to meet an obligation of that amount one month later. If the cost of buying or selling securities is $0.50 per transaction plus $0.10 per $100 of securities bought or sold, it will cost $2.50 ($0.50 + $0.10[20]) to invest the $2,000 in securities at the beginning of the period and an equal amount to sell the securities when the payment needs to be made, or a total cost of $5 for investment and disinvestment. If the interest rate is 4 percent per annum, the interest earnings on $2,000 for one month will amount to $6.67 (0.04[$2,000] ÷ 12), and the net return from investment in the securities will be $1.67 ($6.67 − $5). On the other hand, if the interest rate is 2 percent, the interest earnings will be $3.33 (0.02[$2,000 ÷ 12]), which is less than the $5 cost involved. At an interest rate of 3 percent, the earnings ($5) will be just equal to the cost. Thus, apart from the risks involved, it will pay to invest in securities for the one month period if the interest rate is in excess of 3 percent. For a much more sophisticated analysis of this matter, see W. J. Baumol, "The Transactions Demand for Cash: An Inventory Theoretic," *Quarterly Journal of Economics*, Vol. LXVI, November 1952, pp. 545–56; and James Tobin, "The Interest Elasticity of the Transactions Demand for Cash," *Review of Economics and Statistics*, Vol. XXXVIII, August 1956, pp. 241–47.

these considerations, it is now fairly generally conceded that the demand for transactions balances—especially on the part of larger and more sophisticated spending units—is sensitive to interest rates. We may conclude that the transactions demand for money varies directly with the level of income (GNP) and inversely with the rate of interest.

The Precautionary Motive. Money—currency and demand deposits—has an important advantage over other assets arising out of the fact that it can readily be used to buy goods, services, or securities. Any asset other than money must ordinarily first be sold in a market or otherwise converted into money, which can then be used to purchase the desired item. That is, money is the one asset that possesses perfect liquidity. Of course, money's advantage of perfect liquidity is counterbalanced by the important disadvantage that it pays its holder no appreciable interest or dividend returns.[15]

It seems clear that most spending units hold some amount of cash balances in excess of the minimum called for by the transactions considerations discussed in the previous section.[16] To some extent, cash balances are held for what have come to be called *precautionary purposes.* In part, such balances are held in order to be prepared to *deal with emergencies.* The family going on an extended vacation trip prefers to have some cash in excess of the bare needs of food, lodging, entertainment, and so on, to be used in case some member of the family should become ill and require treatment or in case the family car should break down. Similarly, the business firm ordinarily carries some cash or other relatively liquid assets in excess of bare transactions needs in order to be prepared to deal with emergencies, such as the failure of one of its customers to pay for merchandise at the expected time. The firm would not want to be forced into the position of being unable to meet its own obligations as a result of troubles experienced by its customers.

In addition to the desire to be prepared to deal with unexpected emergencies, precautionary balances may also be held in order to be able to *take advantage of bargains.* For example, it is a great advantage to a family to have sufficient cash on hand to be able to purchase clothing for the children at a time when merchandise goes on sale at reduced prices, even though such

---

[15] Most banks levy service charges against their customers for each check written and for each deposit made. A credit is usually allowed against these service charges which depends upon the size of the customer's average balance. Thus, within limits, an increase in the average balance in a checking account yields a return in the form of a reduction in service charges. In principle, of course, there is no reason why interest could not be paid on checking accounts, and this was, in fact, common practice prior to the mid-1930's. At that time, it was felt that the payment of such interest had encouraged unsound banking practices and contributed to the large number of bank failures that occurred after 1929, and since that time all banking jurisdictions in the United States have prohibited the payment of interest on demand deposits.

[16] For an excellent discussion of the motives for holding money, especially the precautionary motive, see A. G. Hart and P. B. Kenen *Money, Debt, and Economic Activity* (3d ed.; New York: Prentice-Hall, Inc., 1961), chap. xiv.

purchases are not called for by immediate needs. Similarly, a business firm may find that the possession of liquid assets in excess of immediate requirements permits it to take advantage from time to time of unexpected opportunities to make highly advantageous deals.

Thus, precautionary balances are held to meet emergencies and to take advantage of bargains. It is immediately necessary to point out, however, that resources available for these purposes are not limited to cash balances. The spending unit may also hold time deposits which yield an interest return and which can ordinarily be turned into cash on short notice; or it may hold government or other securities which can quickly be sold to raise the needed funds. Furthermore, the precautionary motive may be met at least partially by having an open "line of credit" with a bank—that is, an arrangement under which the spending unit is guaranteed access to bank credit should a need unexpectedly arise, with a moderate fee ordinarily being charged for the service. The development and increasingly widespread use of credit cards in recent years has undoubtedly reduced the precautionary demand for cash balances by households. However, all of the substitutes for money for purposes of satisfying precautionary needs have some disadvantages either in the form of monetary costs (in the case of lines of credit), time required for conversion into cash (in the case of time deposits), or risk of loss due to possible price fluctuations (in the case of most types of securities). On balance, therefore, it is commonly desirable to hold at least a portion of the assets needed to satisfy the precautionary motive in the form of cash balances.

Although it is virtually impossible to isolate cash balances held for precautionary purposes from other cash balances and study separately the factors determining the demand for precautionary balances, theoretical considerations suggest that this demand would be systematically related to the interest rate and the level of income. The carrying of precautionary balances is in part a convenience, the demand for which, like that for any other good or service, presumably depends upon the price. The price in this case is the interest foregone by carrying sterile cash balances rather than interest-bearing claims. A lower rate of interest would presumably encourage the holding of larger precautionary cash balances in two ways: (*a*) by cheapening provisions for liquidity for precautionary purposes generally, and (*b*) by increasing the attractiveness of holding cash for the purpose, as compared with time deposits and marketable securities. Thus, the demand for cash balances for precautionary purposes presumably increases as the interest rate falls.

The demand for precautionary balances can be expected to increase as income rises for two reasons. First, just as persons buy more food, clothing, and other goods and services as their incomes rise, so they can be expected to purchase more of that particular type of convenience and security that can be derived from the holding of larger precautionary balances. Second, as income levels rise, the scale of business operations and of financial commitments increases, thus increasing the "need" for precautionary balances. We may conclude that the *demand for precautionary cash balances varies inversely with the rate of interest and directly with the level of income.* Finally, it should be noted that the precautionary demand for money may at times be *quite sensitive to changes in expectations and the general business outlook.* For example, when fears of depression and unemployment make their appearance, both households and business firms may strive to increase their precautionary balances as a means of strengthening their financial positions so as to be able to ride out the impending storm. On the other hand, when the business outlook improves and optimism begins to prevail, the demand for precautionary balances may be reduced.

The Speculative Motive. Broadly speaking, there are three basic forms in which a person may hold his wealth: money, debt claims, and goods. Ownership of business enterprises in the form, for example, of common stock which represents a proportionate participation in the earnings and assets of a corporation, is referred to as "equity investment" or "equities." Since business enterprises produce goods (or services) for sale, their fortunes are tied to the forces that cause fluctuations in the value of the goods produced; and, consequently, for purposes of our present rather broad classifications, equities may be regarded as having essentially the same investment properties as goods.

Each of these three broad classes of wealth has its own peculiar advantages and disadvantages to the investor. As we have seen, money has the advantage of perfect liquidity combined with the disadvantage of yielding no income in the form of dividends, interest, or capital appreciation. Debt claims yield an interest return, and in the case of marketable debt are subject to price fluctuations which are inverse to movements of interest rates, as explained earlier in this chapter. This is, when interest rates are expected to fall, debt claims may be attractive since there is a prospect of capital gains to supplement interest earnings. On the other hand, when interest rates are expected to rise, debt claims are less attractive since their ownership may entail capital losses. Both money and debt differ from goods (and equities) in the sense that their values do not fluctuate with the price level of goods and services. Thus, both money and debt may be unattractive to investors at times when the price level is expected to rise, since their real value may be expected to fall at such times. Conversely, when the price level is expected to fall, both money and debt may be attractive since appreciation in their real value may be expected. Investment in goods or in equities, on the other hand, is attractive when the price level is expected to rise, since their prices can be expected to move in the same direction as the price level. On the other hand, when the price level is expected to fall, goods or equities are less attractive.

These are, of course, broad generalizations. There are in practice many types of debt claims in terms of maturity, risk, and other contractual provisions. Similarly, there are many goods in which one may invest, as well as a wide variety of equity investments in enterprises of varying degrees of success producing a wide range of goods and services; not all of these are equally attractive investments at any particular time. Success in managing investments is as much a matter of choosing the particular bonds, commodities, or stocks to be purchased or sold at a particular time as it is a matter of timing shifts of funds between the broad categories of money, debt, and goods or equities to take advantage of expected movements of prices and interest rates. Nevertheless, it is this latter aspect of the problem that concerns us at the present time.

In the model we have been developing, the expected returns from holding goods (or investing in facilities for producing goods or services) are reflected in the marginal efficiency of investment schedule. Our present discussion, on the other hand, relates to the factors determining the relative attractiveness of money and debt claims to investors.

Superficially it would appear that once the transactions and precautionary motives for holding money have been taken care of, the investor would always be better off to hold debt claims rather than money, since debt yields interest while money does not. Upon careful examination, however, it is apparent that there may be circumstances in which, strictly from the standpoint of maximizing expected returns on investment, it is preferable to hold money rather than debt.

The problem can best be explored by means of a simple illustration.[17] Suppose an investor has a specified sum of money to be allocated between money and debt claims and that the only debt claims available are 4 percent consols having a face value of $1,000. Let us assume that initially the interest rate on consols is 4 percent, so that consols are selling at par, and that the investor is holding half of his wealth in money form and half in consols. We shall consider the adjustments he might make in the composition of his wealth as the interest rate on consols changes.

The details of the example are presented in Table 11–2. In constructing the example, it was assumed that the investor has a "time horizon" of one year—i.e., that he bases his investment decisions on the returns to be expected from holding consols for one year, taking into account not only contractual interest payments but also any capital gains or losses to be expected during that time. It will become apparent from the example that the investor's behavior will be crucially affected by his interest rate expectations. When the interest rate is at its initial level of 4 percent, it is assumed that the investor expects interest rates to be at that same level of 4 percent at the end of one year, so that there is no expected capital gain or loss to be

[17] For a much more sophisticated analysis of this question, see James Tobin, "Liquidity Preference as Behavior Towards Risk," *Review of Economic Studies,* Vol. XXV, February 1958, pp. 65–86.

experienced from holding a consol for that period, and the expected yield for one year is 4 percent.

Suppose, however, that the interest rate on consols rises to 4.5 percent, so that the price falls to $888.89 ($40 ÷ 0.045).[18] It is assumed in Table 11–2 that the rise in the current interest rate on consols from 4 to 4.5 percent causes the investor to revise his expectations concerning the interest rate at the end of one year upward from 4 to 4.25 percent. Thus, the price expected at the end of a year is now $941.18 ($40 ÷ 0.0425), and the investor will expect a capital gain of $52.29 ($941.18 – $888.89) from investing in

TABLE 11–2

Expected Yields from Holding 4 Percent $1,000 Consol for One Year at Various Levels of Current and Expected Future Interest Rates

Present Interest Rate (Percent)	*Expected Future Interest Rate (Percent)*	*Present Price**	*Expected Future Price†*	*Interest Earnings*	*Capital Gain or Loss (–)‡*	*Total Return§*	*Rate of Return‖ (Percent)*	*Composition of Wealth*
5.00	4.50	$ 800.00	$ 888.89	$40	$ 88.89	$128.89	16.1	All bonds
4.50	4.25	888.89	941.18	40	52.29	92.29	10.4	Mostly bonds
4.00	4.00	1,000.00	1,000.00	40		40.00	4.0	Half bonds
3.50	3.75	1,142.86	1,066.27	40	–76.19	–36.19	–3.2	Few bonds
3.00	3.50	1,333.33	1,142.86	40	–190.47	–150.47	–11.3	No bonds

* $40 ÷ present interest rate.
† $40 ÷ expected future interest rate.
‡ Expected future price – present price.
§ Interest earnings + capital gain or loss.
‖ Total return ÷ present price.

a consol and holding it for one year. When interest of $40 is added in, the total return expected for one year is $92.29; the rate of return for one year is thus 10.4 percent ($92.29 ÷ $888.89).

The other entries in Table 11–2 are calculated in a similar fashion. The general rule that is assumed to govern the investor's interest rate expectations is that any change in the current interest rate on consols causes the investor to adjust his expectations concerning the rate at the end of one year in the same direction as the movement of the current interest rate and by an amount equal to one half of the change in the current interest rate. One interpretation of this rule is that the investor believes any change in the interest rate away from 4 percent will be temporary and that by the

---

[18] This and the succeeding calculations of the price of consols (including those in Table 11–2) employ the formula $V = cP/r$ (see p. 214 above). In this particular calculation, c (the coupon rate) is 0.04, P (the principal value) is $1,000, and r (the market rate of interest) is 0.045, so that $V = (0.04)(\$1{,}000)/0.045 = \$40/0.045 =$ $888.89.

expiration of one year the interest rate will have moved halfway back toward the 4 percent level.

Under these conditions, as can be seen from the next-to-last column of Table 11–2, the return to be expected from holding consols for one year moves up and down with the current interest rate, but the movement is substantially magnified. Thus, the expected return is 16.1 percent when the current interest rate is 5 percent and (as explained above) 10.4 percent when the current interest rate is 4.5 percent. It is especially to be noted that when the current interest rate *falls* to 3.5 percent, the expected yield for one year becomes *negative* to the extent of 3.2 percent, due to the fact that the expected capital loss ($76.19) exceeds the interest earnings ($40). When the current interest rate falls to 3 percent, the expected yield for one year becomes a negative 11.3 percent.

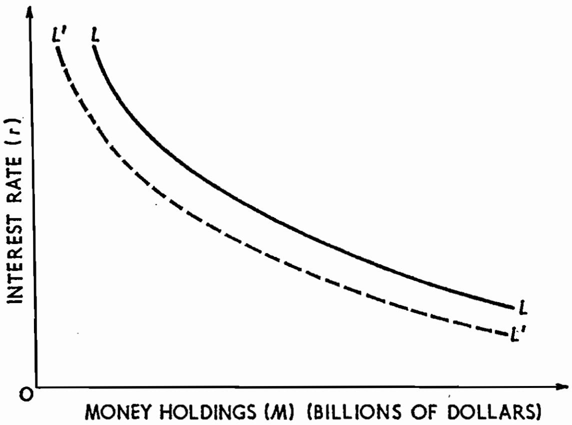
It is now apparent that an investor having expectations such as those assumed in constructing Table 11–2 might under some circumstances prefer to hold noninterest-bearing cash balances rather than interest-bearing securities, because he expects the price of the securities to decline over the period of his calculations by more than his interest earnings and leave him with a net loss.

At first glance, in fact, it would seem that the investor would simply invest all of his resources in bonds if he expected them to yield a positive return and all of his resources in cash if the expected yield on bonds over his horizon was negative. What this overlooks, however, is uncertainty. The investor knows that he may turn out to be wrong in his expectations concerning future changes in the interest rate; consequently, he will not necessarily be prepared to stake all of his resources on his expectations. However, when his expectations are of the type assumed, the expected return from holding bonds rises as the interest rate rises and falls as the interest rate falls; moreover the further the interest rate departs from 4 percent, presumably the greater becomes the degree of certainty with which the investor expects the rate to move back in the direction of 4 percent. Consequently, he will tend to increase the proportion of his resources held in bonds at the expense of the proportion held in cash as the interest rate rises; and, conversely, he will tend to increase his holdings of cash at the expense of bonds as the interest rate declines. This behavior is indicated in the last column of Table 11–2. If investors generally behave in the fashion postulated in this example, the speculative demand for money will take the form of a downward sloping schedule such as that shown in Figure 11–1; that is, the speculative demand for money will increase as the interest rate falls.

The analysis outlined above depicts the typical investor as having a kind of conventional view concerning the interest rate, or, to put it another way, as having in mind a "normal" level of the interest rate toward which he expects the actual interest rate to gravitate. In the example developed in Table 11–2, this conventional or normal interest rate is 4 percent; the further the interest rate departs from this level, the stronger becomes the investor's anticipation that the rate will move back in the direction of the 4 percent level. The conventional or normal interest rate is presumably related to the past experience of investors and reflects the fact that the interest rate is a rather "sticky" price which does not fluctuate over a wide range and which tends to move up and down with changing economic conditions. Of course, events may occur which cause a widespread revision of investors' expectations concerning the conventional or normal interest rate. For example, if the monetary authorities should succeed in convincing investors

FIGURE 11-1

Schedule of the Speculative Demand for Money



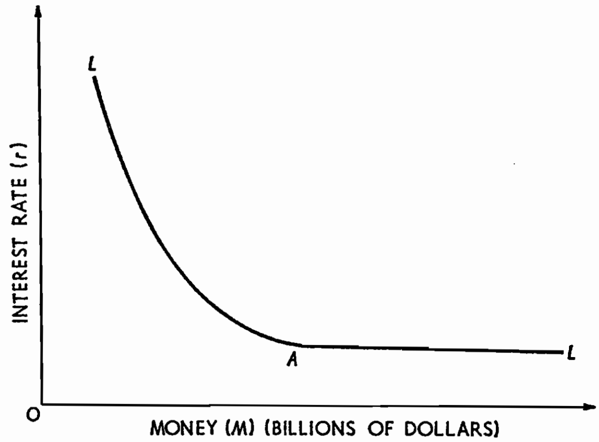
generally of their intention and ability to bring about a permanent decline in interest rates, investors might revise downward their views as to the conventional or normal rate from the neighborhood of 4 percent to that of 3 percent. Such a development would cause the speculative demand curve to shift downward—for example, from position *LL* to position *L′L′* in Figure 11–1.

It is possible that if the interest rate falls sufficiently far below normal, investors will become firmly and unanimously convinced that it must rise in the near future. In this situation, the speculative demand for money may become completely elastic—that is, the schedule may become entirely horizontal as it does to the right of point *A* of the *LL* schedule in Figure 11–2. This situation, often referred to as a *liquidity trap,* means that the speculative demand for money becomes virtually a bottomless pit; that is, investors will absorb unlimited quantities of money into idle balances without using any of it to buy bonds. This possibility, first suggested by Lord Keynes, has come to be an almost standard explanation of the alleged ineffectiveness of monetary policy in time of severe depression. We shall return to this matter later in this chapter.

We may conclude from the above analysis that the speculative demand for money depends upon the interest rate and that the demand increases as the interest rate declines. It is very important to notice, however, that the analysis developed above assumes that the only debt instruments available to the investor are long-term bonds (consols in the example worked out in Table 11–2). However, the theory is very substantially altered when the

FIGURE 11–2

Perfectly Elastic Speculative Demand for Money



existence of short-term debt is taken into account, since in this case the investor has the alternative of investing his funds in interest-bearing short-term debt when long-term bonds become unattractive.[19] After the appear-

[19] Suppose the interest rate on long-term bonds falls below "normal" and the investor expects it to rise in the future so that he will experience a capital loss sufficient to wipe out his interest return over his time horizon. If the only alternative financial asset available to him is noninteresting-bearing money (demand deposits and currency), he may be better off to hold this than bonds so that he will shift into money as in Table 11–2. If, however, as is the case in reality, there are debt instruments available to him which yield an interest return and which will mature and be paid off at their par value within his time horizon, he can invest in these debts without risk of capital loss. Thus, in practice, when interest rates fall in this way, investors have a tendency to shift from long-term debt to short-term debt, causing short-term rates to fall below long-term rates. Conversely, when rates rise above the normal level so that investors expect to reap capital gains from holding long-term bonds, they tend to sell short-term debt and shift into long-term debt, causing short-term interest rates to rise above long-term interest rates. As a consequence, speculative behavior of the kind depicted in Table 11–2 seems to govern movements of the term structure of interest rates rather than being an important factor determining the demand for money. This matter is taken up more extensively in Chapter 14.

ance of Lord Keynes' *General Theory* a quarter of a century ago, the tendency was to regard the demand for money as being sensitive to the interest rate as a result of the operation of the speculative motive. Partly as a result of the complications introduced into the theory due to the existence of short-term debt and partly as a result of research on transactions demand, the recent tendency has been to attribute the greater portion of the interest sensitivity of the demand for money to the transactions and precautionary motives.[20]

The Total Demand for Money

There are various ways of formulating the demand for money which would be consistent with the requirements that the demand increase as income rises and decrease as the interest rate rises. However, we shall explain one very simple formulation which seems to square reasonably well with the facts and also fits in with the above discussion. Suppose we start with the hypothesis that the ratio of money to income is a linear (straight-line) function of the reciprocal of the interest rate. That is,

$$\frac{M}{Y} = a\frac{1}{r} + b \qquad (7)$$

Here M is the money stock, Y is gross national product, r is the interest rate, and a and b are constants to be estimated by statistical procedures.[21]

Using the techniques of regression analysis, it has been found that the function of the type of Equation 7 above that best explains the relationship between M/Y and $1/r$ for the long period from 1892 to 1958 is[22]

---

[20] It should be noted that the speculative demand for money is critically dependent upon the type of expectations that are assumed to prevail. If instead of expecting interest rate movements to be self-reversing, as was assumed in the above example, investors expect rate movements to be cumulative, the nature of the speculative demand for money is drastically altered. For example, if a rise in the interest rate on consols from 4 to 4.5 percent causes an investor to expect the rate to continue to rise to 4.75 percent by the end of the year (instead of expecting the rate to fall back to 4.25 percent, as is assumed in Table 11–2), the expected price at the end of the year will be $842.11, and the investor will expect to incur a capital loss of $46.78 ($888.89 — $842.11), which will more than wipe out the interest of $40. In this case, a rise in the interest rate will encourage him to *sell* bonds and shift into money, thus giving the speculative demand schedule for money a *positive rather than a negative slope.* However, while such expectations may prevail for short periods of time, there is much indirect evidence that the self-reversing type of expectations described in the text and assumed in constructing Table 11–2 are normally prevalent.

[21] This particular formulation of the relationships among money, income, and interest rate was first suggested in H. A. Latane, "Cash Balances and the Interest Rate—A Pragmatic Approach," *Review of Economics and Statistics,* Vol. XXXVI, November 1954, pp. 456–60.

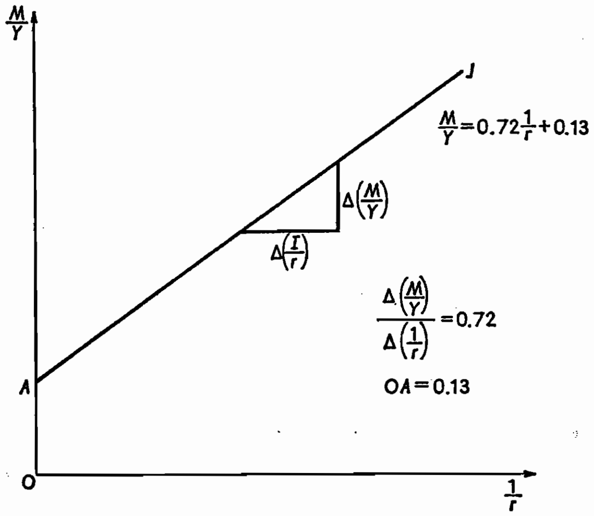
[22] This result is taken from C. F. Christ, "Interest Rates and 'Portfolio Selection' among Liquid Assets," in *Measurement in Economics* (Stanford, Calif.: Stanford University Press, 1963), pp. 201–18.

$$\frac{M}{Y} = 0.72\frac{1}{r} + 0.13 \qquad (8)$$

as depicted graphically in Figure 11–3. The interest rate used here is the yield on long-term high-grade corporate bonds expressed as a percentage. Thus, the values of the constants in Equation 7 which have the greatest explanatory value for the period considered are $a = 0.72$ and $b = 0.13$. This

FIGURE 11–3

Empirical Relation between Ratio of Money to Income and Reciprocal of Interest Rate



relation with $1/r$ "explains" about 75 percent of the variation in M/Y from year to year. The remaining 25 percent that is not explained may be attributed to shifts from time to time in the demand function for money due to factors not included in the equation. For example, during a severe depression such as that of the 1930's, the precautionary demand for money may shift upward as economic units desire to hold more money to protect themselves against the ravages of unemployment than would be expected on the basis of the long-term relationship.

Superficially Equation 8 does not appear to have much similarity to the analysis developed earlier in this chapter; however, it is a simple matter to convert it into a much more familiar form. Multiplying both sides of the equation by Y, we obtain

$$M = 0.72\frac{Y}{r} + 0.13Y \qquad (9)$$

According to the equation in this form, the demand for money increases in proportion to income if the rate of interest is held constant. Thus, if the rate of interest is 4 percent, we have

$$M = 0.72\frac{Y}{4} + 0.13Y = 0.18Y + 0.13Y = 0.31Y$$

That is, at 4 percent interest, the amount of money demanded is 31 percent of GNP. Similarly, at 6 percent interest, the amount of money demanded is 25 percent of GNP.

Suppose now that we hold the GNP constant at $800 billion and see what happens to the demand for money as the interest rate changes. Substituting $800 billion for Y in Equation 9, we obtain

$$M = 0.72\frac{800}{r} + 0.13(800)$$

or

$$M = \frac{576}{r} + 104 \tag{10}$$

Substituting values from 1 percent to 8 percent for r in Equation 10, we obtain the amount of money demanded at each of these interest rates. The results are shown in the second column of the following table:

Interest Rate (r) (Percent)	*Amount of Money Demanded (Billion $)*	
	Y = $800 Billion	*Y = $1,000 Billion*
1	680.0	850.0
2	392.0	490.0
3	296.0	370.0
4	248.0	310.0
5	219.2	274.0
6	200.0	250.0
7	186.3	232.8
8	176.0	220.0

In similar fashion, the amount of money demanded at each interest rate when GNP is $1,000 billion is given by the equation

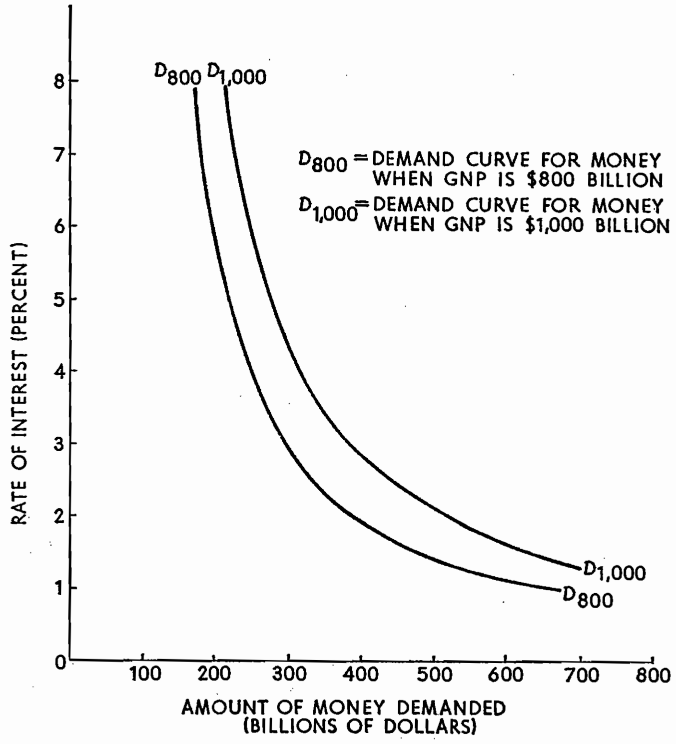
$$M = \frac{720}{r} + 130$$

and the amounts demanded at interest rates from 1 to 8 percent are given in the last column of the above table.

Demand curves for money for GNP of $800 billion and GNP of $1,000 billion are shown in Figure 11–4. These curves, which were obtained by plotting the figures shown in the table above, show the amounts of money that would be demanded at various interest rates for the indicated levels of GNP, assuming the demand for money is of the type described in Equation

FIGURE 11–4

Empirical Demand Curves for Money at GNP Levels of $800 Billion and $1,000 Billion



8 or Equation 9 above. It is customary to describe demand curves for money such as those shown in Figure 11–4 as *liquidity preference* schedules.

In terms of the analysis developed earlier in this chapter, the fact that, as shown in Figure 11–4, the liquidity preference schedule shifts to the right when the GNP rises reflects the fact that the demand for money to satisfy the transactions motive—and probably to some extent the precautionary motive—rises as income rises. The fact that at a given level of GNP the liquidity preference schedule slopes downward to the right probably reflects primarily the fact that the quantity of money demanded to satisfy the transactions motive increases as the interest rate falls, although it may also reflect to some extent the response of the precautionary and speculative demands.

The particular demand-for-money function employed above (Equation 8) is one of several that have been obtained by various investigators, and it was used for illustrative purposes. Accordingly, the student is warned not to take the exact details of this equation too seriously. A full examination of the controversial issues that have been uncovered as a result of recent research on the demand for money would require a detailed analysis of the financial sector of the economy that is well beyond the scope of this book.[23] However, while the various investigations have turned up results that differ in detail, nearly all of the investigations agree on the fundamental proposition that the demand for money increases as income increases and as interest rates decline.[24]

It may be noted that the demand for money is undoubtedly affected by the availability of close substitutes for money and by the general financial structure of the economy. The functioning of the financial system has been changed from time to time as a result of innovations in financing techniques, such as the introduction of the negotiable time certificate of deposit in the 1960's, and such dynamic changes have probably affected the demand for money significantly. Thus, while there is strong evidence that the demand for money depends on income and interest rates, the nature of the relationship has probably changed from time to time as a result of financial innovations.

THE DETERMINATION OF THE INTEREST RATE

We have discussed above the demand for money at considerable length, considering the individual motives for holding money and indicating the nature of the total demand curve for money. We shall now bring the

[23] For a useful and informative survey of recent research on this subject, the student is referred to R. L. Teigen, "The Demand for and Supply of Money" in W. L. Smith and R. L. Teigen (eds.), *Readings in Money, National Income, and Stabilization Policy* (2d ed.; Homewood, Ill.: Richard D. Irwin, Inc., 1970).

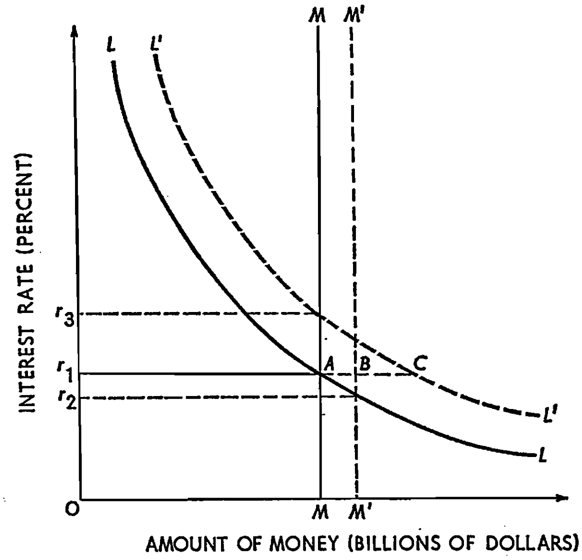
[24] The one major study which reaches significantly different results is that of Milton Friedman ("The Demand for Money: Some Theoretical and Empirical Results," *Journal of Political Economy,* Vol. LXVII, August 1959, pp. 327–51, reprinted in R. A. Gordon and L. R. Klein [eds.], *Readings in Business Cycles* [Homewood, Ill.: Richard D. Irwin, Inc., 1965], pp. 427–55). Friedman defines money to include time and savings deposits in commercial banks (in addition to demand deposits and currency, which are the sole items included in the usual definition), and he postulates that the demand for money depends on "permanent income" (for an explanation of this concept as used by Friedman, see Chapter 5 above, pp. 108–10) rather than current income as measured in the national income accounts. Using these rather unusual concepts and definitions, Friedman succeeds in explaining the demand for money fairly satisfactorily without employing the interest rate as an explanatory variable (although he recognizes that, *in principle* the demand for money should depend on the interest rate). With the exception of Friedman's work, all of the investigations of which the present writer is aware have found that interest rates have a significant effect on the demand for money.

demand for money together with the supply of money to show how the interaction of the two determines the interest rate.

Let us suppose, to begin with, that the level of GNP is given and that the demand for money at that level of GNP is given by the liquidity preference schedule *LL* in Figure 11–5. The supply of money is simply the stock of money available to satisfy the demand. For our present purposes, however, it

FIGURE 11–5

Determination of the Rate of Interest by Demand for and Supply of Money



will suffice to view the stock of money as simply being determined by the central bank or other monetary authority through its ability to control the available supply of cash reserves to the banking system, using for this purpose such devices as open market purchases and sales of government securities or changes in the percentages of their deposit liabilities that commercial banks are required to hold in cash. Thus, the money supply (stock) schedule is a vertical line which shifts to the right when the central bank takes action to increase the money supply and to the left when measures are taken to reduce the money supply. Such a money supply schedule is represented, for example, by the vertical line *MM* in Figure 11–5; let us suppose that this is our money supply schedule to begin with.

If *LL* is the liquidity preference schedule or demand for money and *MM* is the supply curve of money, the rate of interest will be established at r_1 where demand and supply are equal. If the interest rate is higher than r_1, the supply of money will exceed the demand, thus driving down the rate of interest; similarly if the rate is below r_1, the demand will exceed supply, thus driving up the rate of interest.

Suppose now that the central bank takes steps to increase the money supply, causing the supply curve to shift to the right and take up the position *M′M′* in Figure 11–5. At the prevailing interest rate, r_1, the supply of money now exceeds demand by the amount *AB.* The holders of surplus money use it to buy bonds, thus raising the price of bonds and lowering the interest rate. Assuming that the demand curve remains unchanged, the interest rate will continue to fall until it reaches r_2, at which level the demand for money will have increased sufficiently to induce people to hold the expanded stock of money.

The interest rate may also be changed as a result of an autonomous change in the liquidity preference schedule. For example, if expectations concerning future business conditions should deteriorate, the demand for money might increase as spending units attempt to increase their holdings of money for precautionary purposes. Suppose such a development causes the demand curve in Figure 11–5 to shift from *LL* to *L′L′*. At the ruling rate of interest, r_1, the demand for money now exceeds the supply by the amount *AC.* Spending units attempt to achieve the desired increase in their money holdings by selling bonds. Although they cannot obtain more money in the aggregate since the supply remains fixed, their efforts result in a fall in the price of bonds and a rise in the interest rate. Assuming the demand curve remains at the new position *L′L′*, the rate of interest will rise until it reaches r_3, at which point the new demand comes into balance with the fixed supply, *MM.*

The reader should realize that the treatment of the financial sector of the economy in this chapter has been considerably simplified. Two aspects of this simplification are especially worth noting.

1. Throughout this chapter (and also in our discussion of the marginal efficiency of investment in Chapter 8), we have talked about "the interest rate" as though there were only one such rate. In fact, of course, there are many interest rates, which differ according to the maturity of the debt instrument and according to the degree of risk adjudged to attach to that instrument.

2. In drawing Figure 11–5, we assumed that the stock of money could be directly and exactly controlled by the central bank. This justified the treatment of the money supply schedule as a vertical line (*MM* or *M′M′*) which was simply shifted to the left or the right by central bank action. Recent research on the financial sector of the U.S. economy has demonstrated that this is a considerable oversimplification. In particular, there are indications that instead of being completely unresponsive to changes in interest rates, as a vertical schedule would imply, the supply of money may be affected by changes in interest rates.

The oversimplified treatment of the financial sector developed in this chapter is a useful first approximation in the development of an overall explanation of income determination. Accordingly we shall adhere to it in Chapters 12 and 13. However, we will reconsider the financial sector in Chapter 14 in order to show how the introduction of some additional complications affects the analysis developed in Chapters 12 and 13.

Chapter 12 MACROECONOMIC EQUILIBRIUM

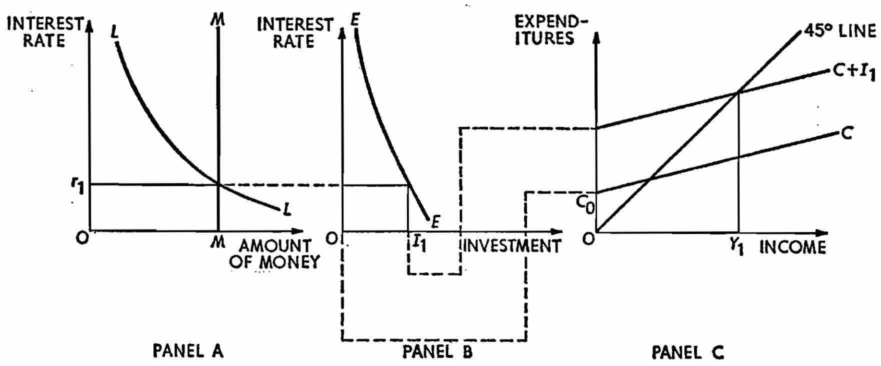
The discussion in Chapter 11 of the effects on the interest rate of changes in the supply of and demand for money considers only the reactions of the money market in isolation from the rest of the economy. For example, in considering the effects of an increase in the money supply it was assumed that the demand for money remained fixed in its original position *LL* in Figure 11–5. While this may be appropriate as long as the analysis relates to the very short run, it is a seriously incomplete description of the economic adjustments involved. The reason for this is that if an increase in the money stock lowers the rate of interest, investment will be stimulated to the extent that the marginal efficiency of investment schedule is elastic, as indicated in Chapter 8. The rise in investment will set off an expansion of income (GNP) via the multiplier mechanism explained in Chapter 6. Since the demand for money for transactions and precautionary purposes increases as income rises, the liquidity preference schedule will not remain in the position *LL* but will shift to the right as income rises. Similarly, in the case of an increase in the demand for money from *LL to L′L′*, the resulting rise in the interest rate will reduce investment, thus causing GNP to decline and inducing a secondary shift in the demand for money back to the left. In order to deal satisfactorily with these indirect repercussions, it is necessary to pull together the main elements of the analysis developed in the earlier chapters of Part II.

A SIMPLIFIED PRESENTATION

A useful, although not entirely satisfactory, way of accomplishing this integration is depicted in Figure 12–1. Panel A on the left of Figure 12–1 shows the liquidity preference schedule (*LL*), i.e., the demand schedule for money, and the externally determined supply of money (*MM*). These two schedules determine the rate of interest, which in this instance is r_1. Turning next to Panel B, the interest rate, determined in Panel A, is brought together with the marginal efficiency of investment schedule (*EE*) to determine the rate of investment I_1, in accordance with the analysis developed in Chapter 8. Panel C presents the consumption function/multiplier analysis, as developed in Chapters 5 and 6. When the amount of investment determined in Panel B is superimposed on the consumption function (C_0C), the intersection of the resulting total expenditure function (labeled $C + I_1$) and the 45° guideline determines the level of income.

FIGURE 12–1

Simplified Presentation of Income Determination



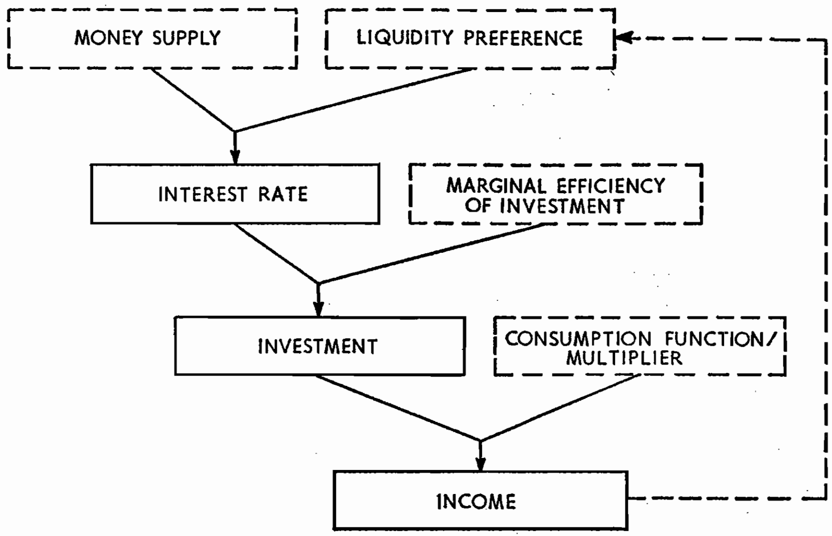
The chain of causation is shown schematically in Figure 12–2 and may be stated briefly in three steps:

1. The money supply and liquidity preference determine the interest rate.
2. The interest rate and the marginal efficiency of investment determine investment.
3. Investment and the consumption function/multiplier mechanism determine income.

While this simple view of causation is useful for the student to get in mind, it does represent an oversimplification of the model we are working with. The reason is that it overlooks the fact that the demand for money for transactions and precautionary purposes depends on income; that is, there is a "feedback" from income to the liquidity preference schedule, which is shown by the long arrow pointing upward to the right in Figure 12–2. That is, in Figure 12–1, it is really not proper to take the position of the *LL* schedule as given in Panel A, since this position depends upon the level of income shown in Panel C.

FIGURE 12–2

Schematic Summary of Factors Determining Income and Related Variables



RECOGNITION OF INTERDEPENDENCE

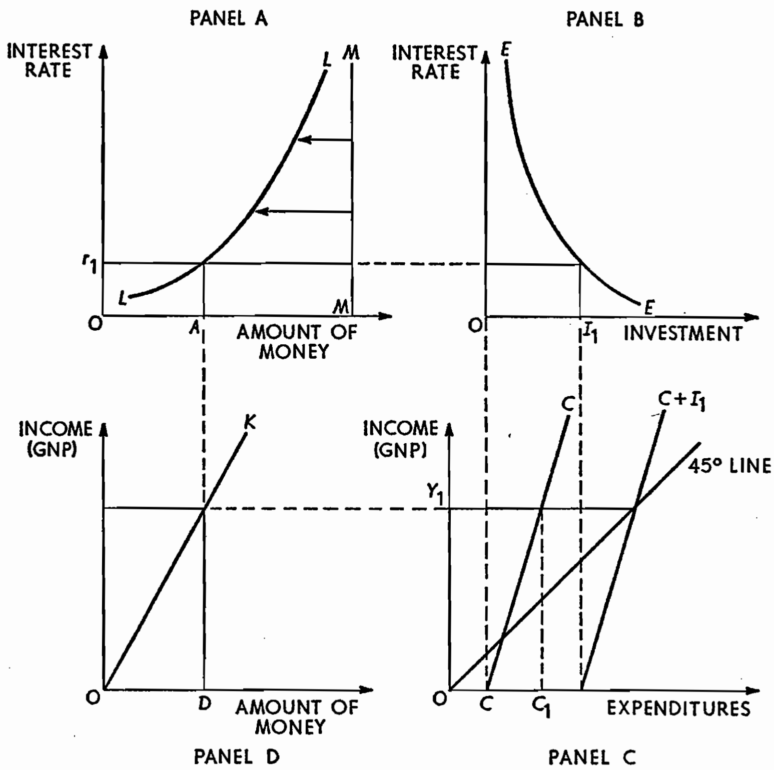
In truth, the model we are using is a *static interdependent* model in which the equilibrium values of all of the variables are *simultaneously* determined. A graphical presentation which shows this interdependence is given in Figure 12–3.[1] This figure should be studied in clockwise fashion beginning with Panel A in the upper left-hand corner. Panel A shows the relation between the interest rate (r) and money (M). It is constructed by measuring a distance *to the right* from the origin equal to the independently determined money supply and erecting a vertical line (MM). In order to facilitate the graphical presentation, the demand for money is assumed to be divisible into two distinct parts, one depending on income and the other depending on the interest rate. For purposes of exposition, we shall refer to the portion dependent on income as the *transactions demand* and the portion dependent on the interest rate as the *asset demand.*[2] In Panel A, the asset

[1] This graphical presentation is a slight modification of the presentation given in I. O. Scott, Jr., "An Exposition of the Keynesian System," *Review of Economic Studies,* Vol. XIX (1), 1950–51, pp. 12–18.

[2] The reader will note that this division of the demand for money is not strictly consistent with the results of the analysis presented earlier in this chapter, since that analysis led us to conclude that both the transactions and precautionary components of demand may depend on income, while the transactions, precautionary, and speculative components of demand may all depend upon the interest rate. However, the inconsistency is chiefly a matter of terminology and has no important effects on our conclusions.

FIGURE 12–3

Graphical Presentation of Macroeconomic Equilibrium



demand schedule (*LL*) is drawn *measuring to the left from the line MM.* That is, at any interest rate the amount of money that would be desired as an asset is given by the distance between the curve *LL* and the line *MM.* Panel B shows the marginal efficiency of investment schedule (*EE*) relating the interest rate to investment. Panel C is simply the customary consumption function/multiplier diagram employing a 45° guideline, except that the axes are reversed as compared with our previous presentation (i.e., income is measured vertically and expenditures horizontally) and the whole diagram is shifted to the left so that the intersection of the consumption function with the horizontal axis falls directly beneath the origin of Panel B. Panel D shows the relation between income and the transactions demand for money —a relation which is assumed to be proportional, as described by the line *OK.*

To see how Figure 12–3 works, suppose we select an interest rate, r_1. Panel A shows that at this interest rate, the amount of money held as an asset would be AM. Tracing over to Panel B, investment would be I_1; carrying this down to the horizontal axis of Panel C and from that point constructing a line (labeled $C + I_1$) parallel to the consumption function, the intersection of this line with the 45° guideline gives the income Y_1, as shown on the vertical axis. Consumption is C_1 as determined by the consumption function (the line CC) at income Y_1 in Panel C. Tracing over to Panel D at income Y_1, necessary transactions balances are found to be equal to OD, which is equal to OA in Panel A.

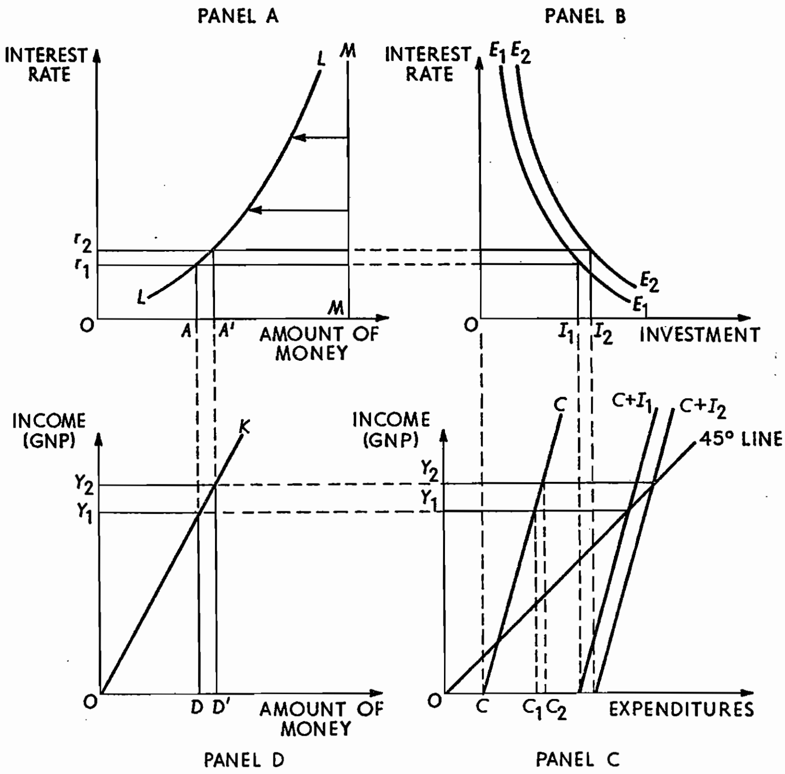
The interest rate, r_1, and the corresponding values of other variables, I_1, Y_1, and C_1, are equilibrium values, as can be seen from the fact that the amount of money held as an asset at interest rate r_1 (the distance AM in Panel A) plus the amount of money held for transactions purposes at income Y_1 (the distance OA in Panel A) is just equal to the existing stock of money (the distance OM in Panel A). The student can verify for himself the fact that if we start with any interest rate other than r_1 we will not come out right. (The solution—i.e., the combination r_1, I_1, and Y_1 which does come out right—was, of course, found by trial and error.) Thus, we have an interdependent system in the sense that there is only one set of values of the variables (r_1, I_1, C_1, Y_1) which simultaneously satisfies all of the relationships.

CHANGES IN CONDITIONS OF MACROECONOMIC EQUILIBRIUM

A situation such as that depicted in Figure 12–3 in which all of the variables in our model—income, consumption, investment, and the rate of interest—have adjusted to their appropriate values in relation to a given supply of money may be described as a position of *macroeconomic equilibrium.* If an autonomous shift occurs in one or more of the schedules shown in Figure 12–3, the equilibrium is upset and the values of all of the variables —income (Y), interest rate (r), investment (I), and consumption (C)— will be changed. This is demonstrated in Figure 12–4, which shows the effects of a rise (i.e., a shift to the right) of the marginal efficiency of investment schedule, which might be the result, for example, of technological change. When the marginal efficiency of investment schedule is in its original position, E_1E_1, the interest rate is r_1, investment I_1, consumption C_1 and income Y_1. When the schedule shifts to E_2E_2, the interest rate rises ro r_2, investment rises to I_2, consumption rises to C_2, and income rises to Y_2. The increase in investment resulting directly from the shift raises the income level via the multiplier. As income rises, the transactions demand for money rises (shown as a movement along the line OK in Panel D). With the money supply fixed, money is drawn out of asset balances (which decline from AM to $A'M'$), and to accomplish this withdrawal the interest rate must rise (from r_1 to r_2). This rise in the interest rate causes a backward movement along the new marginal efficiency of investment schedule

FIGURE 12–4

Effects of a Shift in the Marginal Efficiency of Investment Schedule



(E_2E_2), so that investment does not increase by the full amount of the shift in the *MEI* schedule. That is, with a fixed money supply, the multiplier effects of the rise in investment are somewhat reduced due to the fact that the rise in income increases the demand for money and causes the interest rate to rise.

According to this analysis, changes in the GNP, together with the related variables interest rate, investment, and consumption, may be due to the following four proximate causes:

1. ***Changes in the Money Supply.*** Such changes appear as shifts in the *MM* and *LL* schedules in Figure 12–3 and are the result of actions taken by the central bank or other monetary authority. An increase in the money supply causes both the line *MM* and the curve *LL* to shift to the right by equal amounts, since the *LL* curve is drawn with *MM* as a base, as explained above. Similarly, a decrease in money supply causes *MM* and *LL* to shift to the left.

2. Changes in Liquidity Preference. Such changes appear as shifts (to the left for an increase in liquidity preference and to the right for a decrease) in the *LL* schedule in Panel A and/or shifts (clockwise for an increase in liquidity preference and counterclockwise for a decrease) in the line *OK* in Panel D. Shifts may be due to changes in payment patterns which cause changes in the transactions demand for money, changes in business confidence which cause changes in the precautionary demand, or revisions of the views of investors concerning the "normal" level of the interest rate which cause changes in the speculative demand for money. (See discussion of these matters in Chapter 11.)

3. Changes in the Marginal Efficiency of Investment. Such changes are reflected in shifts (to the right for an increase, to the left for a decrease) in the *EE* schedule as shown in Panel B of Figure 12–3, and an illustration is given in Figure 12–4. These shifts may be due to technological developments, population changes, autonomous shifts in business expectations, and so on (see the discussion in Chapter 9). Government purchases of goods and services may be included, along with private investment, in the *EE* schedule. Thus, an increase in government purchases of goods and services produces a rightward shift of the *EE* schedule, and a decrease in such expenditures produces a leftward shift.

4. Changes in the Consumption Function. Changes in the consumption function are reflected in shifts (to the right for an increase, to the left for a decrease) in the CC_0 schedule in Panel C of Figure 12–3. Changes may be due to such factors as changes in consumer expectations concerning prices or incomes, changes in consumer holdings of liquid assets, and so on. (See the discussion in Chapter 5.) Changes in legislation relating to taxes and transfer payments may also have the effect of shifting the consumption function by changing disposable personal income for a given level of GNP. On the other hand, changes in taxes and transfer payments *resulting from* changes in GNP with given legislation and regulations are reflected in the *slope* of the CC_0 line (i.e., the marginal propensity to consume out of GNP) and thereby in the multiplier. (See the discussion of fiscal variables in Chapter 6.)

These four factors—money supply, liquidity preference, marginal efficiency of investment, and the consumption function—are the basic determinants of GNP and related variables. Accordingly, in order to distinguish them from the variables of the system, the boxes containing them are surrounded by heavy borders in Figure 12–2. No change in GNP can occur unless changes in one or more of these four factors take place. With the model used in constructing Figure 12–3, the *direction* of the effects of changes in each of the four factors on the variables *Y, r, I,* and *C* can be deduced unambiguously and are summarized in Table 12–1. For example, the first line of the table indicates that an increase in the money supply will appear as a shift to the right of the *MM* and *LL* schedules in Panel A of

TABLE 12–1

Effects Produced by Shifts in Various Schedules in Figure 12–3

			Effect on			
Change in:	*Direction of Change*	*How Shown in Figure 12–3*	*r*	*I*	*C*	*Y*
Money supply	Increase	*MM* and *LL* shift to right	Decrease	Increase	Increase	Increase
	Decrease	*MM* and *LL* shift to left	Increase	Decrease	Decrease	Decrease
Liquidity preference	Increase	*LL* shifts to left and/or *OK* revolves clockwise	Increase	Decrease	Decrease	Decrease
	Decrease	*LL* shifts to right and/or *OK* revolves counterclockwise	Decrease	Increase	Increase	Increase
Marginal efficiency of investment	Increase	*MEI* shifts to right	Increase	Increase	Increase	Increase
	Decrease	*MEI* shifts to left	Decrease	Decrease	Decrease	Decrease
Consumption function	Increase	*CC* shifts to right	Increase	Decrease	Increase	Increase
	Decrease	*CC* shifts to left	Decrease	Increase	Decrease	Decrease

Figure 12–3 and will decrease the interest rate and increase investment, consumption, and GNP. The student should analyze for himself the effects of each of the changes by reference to Figure 12–3 and satisfy himself that the effects described in Table 12–1 are correct as to direction. Of course, the *magnitudes of the effects* depend upon the exact shapes and positions of the schedule shown in Figure 12–3—i.e., the *LL* schedule, the *EE* schedule, the *CC* schedule, and the *OK* schedule.

As pointed out in Chapter 5, it is possible that the amount saved (and therefore, of course, the amount spent on consumption) at a given level of income may be affected by the rate of interest. Such an effect, if present, can be taken into account by including in the *EE* curve of Panel B of Figures 12–3 and 12–4 not only the interest sensitivity of investment but the interest sensitivity of consumption as well.[3] In addition, some government expenditures—such as expenditures by state and local governments on schools and highways financed by means of bond issues—may be affected by interest rates, and these may also be included in the *EE* schedule. In fact, the *EE* schedule may be drawn to reflect the interest sensitivity of all expenditures and accordingly may be called the *expenditure schedule* rather than merely the marginal efficiency of investment schedule.

EFFECTS OF A CHANGE IN THE MONEY SUPPLY

The strength of the effect on the level of national income of a change in the money supply depends upon several factors which we shall now take up.

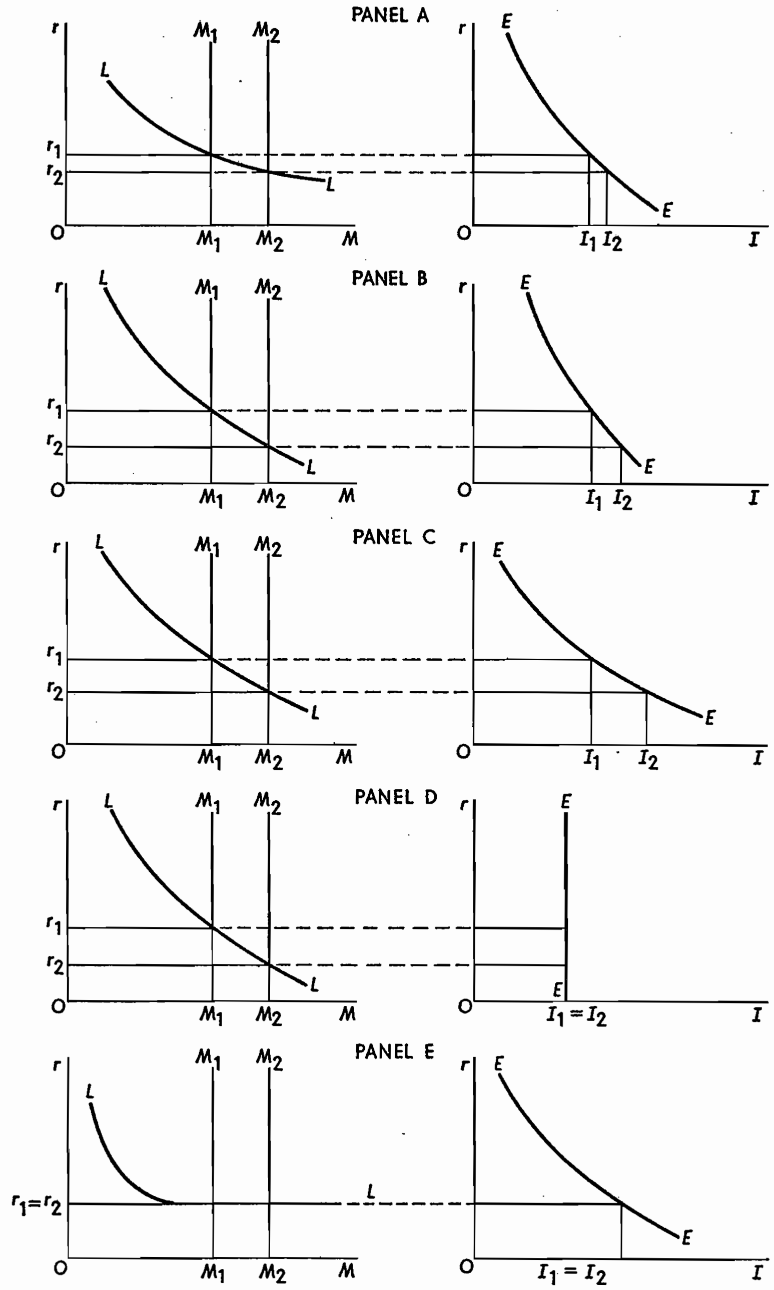
1. ***The Interest Elasticity of Liquidity Preference.*** An elastic liquidity preference schedule is one which is relatively flat, such as the *LL* curve in Panel A of Figure 12–5, so that a small decline in the interest rate causes a large increase in the quantity of money that spending units desire to hold. Other things equal, a change in the money supply will have a greater effect the less elastic, or the more inelastic, is liquidity preference. This can be seen from a comparison between Panels A and B of Figure 12–5. In these two panels the expenditure schedules (*EE*) are exactly the same, while the liquidity preference schedule (*LL*) is less elastic in Panel B than in Panel A. Thus, a given increase in the money supply (from M_1 to M_2) causes a larger decline in the interest rate and therefore causes a larger increase in investment in Panel B than in Panel A, thereby exerting a larger impact on income.

2. ***The Interest Elasticity of Expenditures.*** An elastic expenditure schedule is one which has a gentle slope, so that a small decline in the

[3] As pointed out in Chapter 5, it is theoretically possible for a change in the rate of interest to affect saving (and therefore consumption) in either direction. If such interest rate effects on saving are important—and particularly if saving should decline as the interest rate rises—some of the directions of effects shown in Table 12–1 might be reversed. Since, however, there is little evidence of significant interest rate effects on saving, no effort has been made to take these possibilities into account.

FIGURE 12–5

Examples Showing Effects of Varying Interest Elasticities of the Liquidity Preference and Expenditure Schedules



interest rate causes a substantial increase in expenditures. Other things equal, a change in the money supply will have a greater effect on income the more elastic the expenditure schedule. This can be seen from a comparison of Panels B and C of Figure 12–5. The liquidity preference schedules are exactly the same in Panels B and C, so that an increase in the money supply from M_1 to M_2 causes the interest rate to fall by the same amount. However, expenditures increase to a greater extent in Panel C than in Panel B, because the elasticity of the expenditure schedule is greater in the latter case.

*3. **The Size of the Multiplier.*** The effect of a change in the money supply on income will be larger, other things equal, the larger is the multiplier, or, to put it another way, the larger is the marginal propensity to consume out of GNP. This is because it is through the multiplier that the initial change in expenditures caused by the change in interest rates is converted through the process of respending into a magnified effect on income.

*4. **The Size of the Income Effect on the Demand for Money.*** According to our analysis earlier in this chapter, the demand for money for transactions and precautionary purposes depends on the level of income. Thus, for example, as the level of income rises, the amount of money demanded for transactions and precautionary purposes increases. This increase in the demand for money will have a tendency to raise the rate of interest, thus counteracting part of the decline in the interest rate caused initially by the increase in the money supply. Thus, the stronger the effect of an increase in income on the demand for money, the weaker will be the effect of a change in the money supply on the level of income—that is, if the demand for money increases by 30 percent of the rise in income, the effects will be weaker than if the demand for money changed by only 20 percent of the change in income.

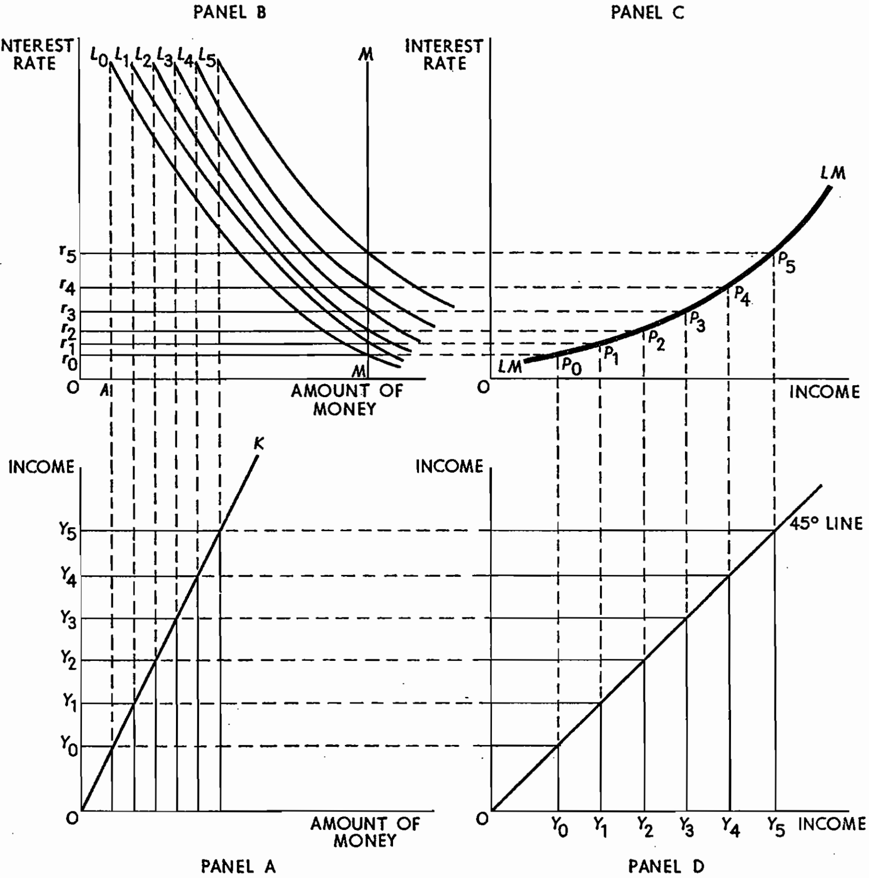
We may summarize as follows: the effect of a change in the money supply on the level of income will be greater (*a*) the less elastic is the liquidity preference schedule, (*b*) the more elastic is the expenditure schedule, (*c*) the larger is the multiplier, and (*d*) the smaller is the income effect on the demand for money. Panels D and E of Figure 12–5 illustrate two possible situations in which a change in the supply of money will have no effect whatever on income. In Panel D, the liquidity preference schedule is rather elastic, so that the increase in money supply from M_1 to M_2 causes a substantial decline in the interest rate; but the expenditure schedule is completely inelastic, so that the fall in the interest rate has no effect on expenditures. Panel E illustrates the so-called "liquidity trap" case in which the liquidity preference schedule is completely elastic, so that the increase in the money supply has no effect on the interest rate—i.e., all of the increase in money supply is absorbed into idle balances—so that even though the expenditure schedule is fairly elastic, there is no increase in expenditures because the interest rate does not fall.

A CONCISE DIAGRAMMATIC SUMMARY OF MACROECONOMIC EQUILIBRIUM

The equilibrium of the money market is brought together and summarized in Figure 12–6. Again, as in the preceding section of this chapter, the demand for money is divided into two parts, one (which we shall call the transactions demand) dependent on income and the other (which we shall call the asset demand) dependent on the interest rate. Panel A in the lower left-hand corner shows the relation between income and the transactions demand for money; this relation is depicted by the line *OK*. Income is measured on the vertical axis and money holdings on the horizontal axis. If we select an income, say Y_0, and trace over horizontally to the line *OK*,

FIGURE 12–6

Graphical Presentation of Equilibrium of the Money Market



the horizontal distance represents the transactions demand for money at that income. We then trace upward to Panel B, which shows the interest rate on the vertical axis and money on the horizontal axis, and draw the asset demand curve for money, labeled L_0L_0, which shows the total amount of money demanded (transactions demand plus asset demand) at each interest rate when income is Y_0. The vertical line *MM* represents the stock of money, which is assumed to be constant; and the intersection between the demand schedule for money, L_0L_0 and the supply of money schedule, *MM,* gives us the interest rate, r_0, which is the equilibrium interest rate corresponding to income Y_0 and money stock *MM, OA* being the transactions demand and *AM* the asset demand. Panel D in the lower right-hand corner shows income on both axes with a 45° line between and merely serves to convert income from a vertical distance in Panel A into a horizontal distance as needed in Panel C. Tracing across at interest rate r_0 from Panel B to Panel C and tracing across at income Y_0 from Panel A to Panel D and then up to Panel C, we obtain the point P_0 in Panel C, whose vertical distance is r_0 and whose horizontal distance is Y_0. This procedure can be repeated starting with various income levels in Panel A (such as Y_1, Y_2, Y_3, Y_4, and Y_5), obtaining in each case the corresponding interest rate (r_1, r_2, r_3, r_4, and r_5). These combinations of income and interest rate at which the money market is equilibrated are shown as a series of points in Panel C (P_1, P_2, P_3, P_4, and P_5). A curve is drawn through these points and labeled *LM* in Panel C.

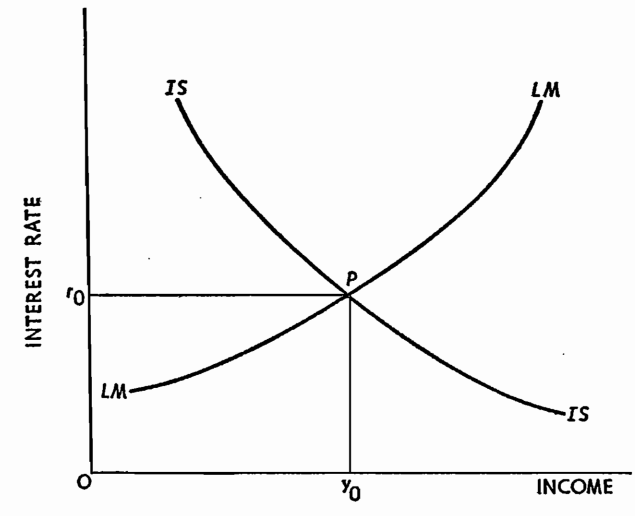
The *LM* curve shows the various combinations of income and interest rate at which the money market will be in equilibrium with a given money supply. The curve slopes upward to the right, because as income rises the amount of money needed to support transactions rises, pushing the total demand for money (in Panel B) to the right and thereby (since the total money supply is fixed) reducing the amount of money remaining to satisfy the asset demand and necessitating a rise in the interest rate.

The *LM* curve can now be brought together with the *IS* curve, which was explained at the end of Chapter 10 and shown in Figure 10–4, to provide a concise summary of the conditions of macroeconomic equilibrium.[4] The result of such a merger of the *IS* and *LM* curves is presented in Figure 12–7. In this diagram, the *IS* curve, as explained earlier, represents all combinations of income and the interest rate for which planned saving equals planned investment or the market for goods and services is in equilibrium; that is, this curve incorporates the marginal efficiency of investment (or expenditure) schedule and the consumption function/multiplier mechanism. The *LM* curve, on the other hand, is the locus of all combinations of

[4] For a further discussion of the *IS* and *LM* curves and their use, see J. R. Hicks, "Mr. Keynes and the 'Classics': A Suggested Interpretation," *Econometrica,* Vol. V, April 1937, pp. 147–159, reprinted in William Fellner and B. F. Haley (eds.), *Readings in the Theory of Income Distribution* (Philadelphia: Blakiston Co., 1946), pp. 461–76; and A. H. Hansen, *Monetary Theory and Fiscal Policy* (New York: McGraw-Hill Book Co., 1949), chap. 5.

FIGURE 12–7

Graphical Summary of Conditions of Macroeconomic Equilibrium



income and the interest rate for which the money market is in equilibrium with a given supply of money. Thus, the point P at which the IS and LM curves intersect represents the combination of income and interest rate at which both the market for goods and services and the money market are in equilibrium and therefore constitutes the monetary equilibrium position of the system.[5]

[5] Actually, in the model we have developed there are three sets of markets: (1) the market for goods and services, (2) the market for bonds, and (3) the market for money. However, in such an economy, for example, if a person's sales of goods and services exceed his purchases of goods and services by $5,000, while his purchase of bonds exceed his sales of bonds by $3,000, his receipts of cash must necessarily exceed his disbursements of cash by $2,000 (i.e., his cash balance must increase by $2,000). If we use the term excess demand to mean the amount (in money terms) by which demand exceeds supply (excess demand being negative if supply exceeds demand), it is necessarily true under all circumstances that

Excess demand for goods and services
+ Excess demand for bonds + Excess demand
for money = 0

(That is, in the above example, − $5,000 + $3,000 + $2,000 = 0.) According to this principle (known as Walras' Law after Leon Walras, a famous 19th-century economic theorist) if two of the three sets of markets are equilibrated—that is, if supply equals demand in these markets, so that excess demand equals zero—it follows that the third set of markets must necessarily be equilibrated also. Thus, if we find the conditions under which two of the sets of markets are equilibrated, the third set will take care of itself. In our analysis, we have chosen to study the conditions of equilibrium in the market for goods and services and in the market for money, and it is therefore not necessary for us to worry explicitly about the market for bonds. For a highly sophisticated analysis of these matters, see Don Patinkin, *Money, Interest, and Prices* (2d ed.; New York: Harper & Row, 1965).

As indicated earlier in this chapter, changes in the level of income and the other related variables (interest rate, investment, and consumption) may be caused by changes in (*a*) the money supply, (*b*) liquidity preference, (*c*) marginal efficiency of investment (or, more broadly, the expenditure schedule), and (*d*) the consumption function. All of these factors are incorporated in Figure 12–7. As was explained in Chapter 10, a shift to the right of the marginal efficiency of investment (or expenditure) schedule as shown in Panel A of Figure 10–4 will cause the *IS* curve to shift to the right, while a shift to the left of the *MEI* schedule will cause the *IS* curve to shift to the left. A rise in the consumption function, as shown (line C_0C) in Panel C of Figure 10–4, will cause the *IS* curve to shift to the right, while a fall in the consumption function will cause the *IS* curve to shift to the left. As indicated above, an increase in the money supply as shown by a shift to the right of the line *MM* in Panel B of Figure 12–6 will cause the *LM* curve to shift to the right, while a decrease in the money supply will cause the *LM* curve to shift to the left. An increase in liquidity preference as shown by a shift to the right of all of the *LL* schedules corresponding to various income levels in Panel B of Figure 12–6, or by a clockwise rotation of the line *OK* in Panel A of Figure 12–6, will cause the *LM* curve to shift to the left, while a fall in liquidity preference will cause the *LM* curve to shift to the right. Thus, all of the factors affecting the level of income that we have discussed are incorporated in the two curves shown in Figure 12–7.

Chapter 13

FURTHER ANALYSIS OF AN ECONOMY WITH A SIMPLE MONETARY SECTOR

In order to illustrate more concretely some of the ideas summarized and explained in Chapter 12, it will be useful to explore the working of a relatively simple model containing both a real sector and a financial sector. In order to keep the algebra as simple as possible, we will employ a model that consists entirely of linear relationships. It should be understood that this is done largely as a matter of convenience and that many of the propositions developed in this chapter would hold for nonlinear models also.

A STATIC LINEAR MODEL OF THE REAL AND FINANCIAL SECTORS

The basic model we shall work with consists of the following eight equations.

$$C = cY_d + C_0 \quad (1)$$
$$I = iY_d - vr + I_0 \quad (2)$$
$$X = xY + X^* \quad (3)$$
$$Y_d = Y - X \quad (4)$$
$$Y = C + I + G^* \quad (5)$$
$$M_d = kY - mr + L_0 \quad (6)$$
$$M_s = M^* \quad (7)$$
$$M_d = M_s \quad (8)$$

The model contains eight endogenous variables, defined as follows: C = personal consumption expenditures; Y_d = disposable income; I = gross private domestic investment; r = the interest rate; X = net taxes (taxes less government transfer payments); Y = gross national product; M_d = the quantity of money (demand deposits and currency) demanded; and M_s = the money supply. There are three exogenous policy parameters which the government may adjust to influence the economy: M^*, the money supply; X^*, the level of the net tax function; and G^*, government purchases of goods and services. There are two other magnitudes to which we shall have occasion to refer, the budget surplus or deficit (B) and private saving (S), which are defined as follows:

$$B = X - G$$
$$S = Y_d - C$$

Equation 1 is a standard linear short-run consumption function of the type depicted in Figure 5–1. Equation 2 is the investment function, which makes investment depend on the interest rate and the *level* of disposable income. It should be emphasized that this equation does *not* contain an investment accelerator; the acceleration principle (whether in the simple or the more flexible stock adjustment form) makes investment depend on the *rate of change* of income. Since the model we are using is static and the acceleration theory is inherently dynamic, that theory cannot be incorporated in this model. The justification for the inclusion of disposable income in the investment function is that it reflects changes in profits, which may affect investment in two ways: (1) higher current profits may create the expectation of higher profits in the future thereby strengthening the incentive to invest, and (2) an increase in profits means an increased flow of internal funds with which to finance investment.

Equation 3 is the tax function, which shows total net taxes (taxes minus transfer payments) as a function of GNP. Equation 4 simply defines disposable income as GNP minus net taxes. The treatment of the deductions to be made from GNP to arrive at disposable income is greatly oversimplified. All forms of taxes are lumped together, and no distinction is made between personal disposable income and aftertax corporate profits; consequently private saving (S) includes both personal and corporate saving. It would be possible to incorporate in this model a more elaborate treatment of corporate profits, dividends, and different types of taxes and transfer payments in accordance with the treatment developed in Chapter 6.[1] However, this would greatly complicate the algebra while adding relatively little to the substance of the analysis.

Equation 5 rounds out the model of the real sector of the economy. It represents the equilibrium condition that total production (Y) must be equal to aggregate demand for final output ($C + I + G^*$), or that aggregate supply must equal aggregate demand.

The last three equations of the model describe a very simple monetary or financial sector of the economy. Equation 6 indicates that the demand for money increases as income increases and declines as the interest rate increases, which accords generally with the discussion of the demand for money in Chapter 11. Equation 7 says that the supply of money is equal to M^*, a magnitude that is assumed to be controlled by the central bank. Equation 8 is the equilibrium condition for the money market, indicating that the demand for money must be equal to the supply of money.

[1] See Chapter 6, pp. 126–33.

It may be noted that all of the slope parameters of the model (c, i, v, x, k, and m) are assumed to be positive. The fact that both investment and the demand for money decline as the interest rate increases is reflected in the presence of minus signs in front of v in Equation 2 and m in Equation 6.

Let us begin by deriving the *IS* and *LM* curves for this model. Substituting Equations 1, 2, 3, and 4 into Equation 5, we obtain

$$Y = c(Y - xY - X^*) + C_0 + i(Y - xY - X^*) - vr + I_0 + G^*$$

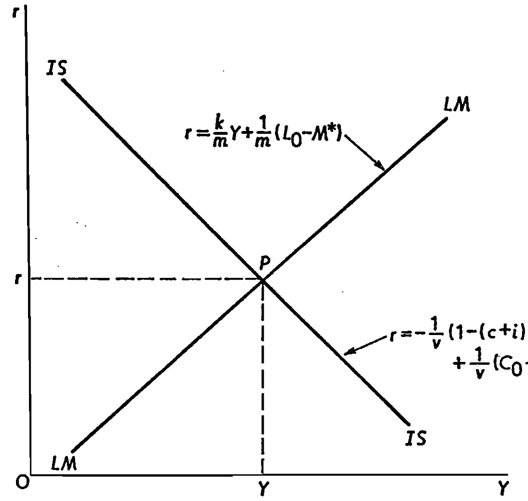
This equation contains only two endogenous variables, Y and r. Solving for r in terms of Y, we obtain

$$r = -\frac{1}{v}(1 - (c + i)(1 - x))Y + \frac{1}{v}(C_0 + I_0 - (c + i)X^* + G^*) \quad (9)$$

This is the *equation of the IS curve.* The expression $(c + i)(1 - x)$ is the

FIGURE 13–1

IS and *LM* Curves for Linear Model Represented by Equations 1–8



marginal propensity to spend on consumption and investment out of GNP. If we assume that this propensity is less than unity, the coefficient of Y, which is $-\frac{1}{v}(1 - (c + i)(1 - x))$, will be negative, so that the *IS* curve will have a negative slope, as shown in Figure 13–1. It may be noted that in the graphical derivation of the *IS* curve that was presented in Figure 10–4, no allowance was made for the effect of income on investment—that is, it was assumed that investment depended only on the interest rate and not on the level of income. While the presence of income in the investment equation would have made it impossible to use the graphical technique of Figure 10–4 to derive the *IS* curve, it is apparent from Equation 9 that a very similar *IS* curve exists for a model such as the present one which contains an income effect on investment. As in the simpler model, the *IS* curve represents all combinations of income and the interest rate which result in equilibrium in the market for goods and services.

Substituting Equations 7 and 8 into Equation 6 we obtain the equation

$$M^* = kY - mr + L_0$$

This equation contains only two endogenous variables, Y and r. Solving for r in terms of Y, we obtain

$$r = \frac{k}{m}Y + \frac{1}{m}(L_0 - M^*) \qquad (10)$$

This is the equation of the *LM* curve. Since both k and m are positive, the equation has a positive slope as indicated in Figure 13–1. Equation 10 and the corresponding curve (*LM*) in Figure 13–1 represent the locus of values of Y and r at which the money market is in equilibrium with a given stock of money (M^*).

Since the real sector of the economy is in equilibrium at all points on the *IS* curve and the financial sector is in equilibrium at all points on the *LM* curve, the point (P in Figure 13–1) of intersection of the two curves is a position of equilibrium for the whole economy. We can find the equilibrium value of Y by eliminating r between Equations 9 and 10. To do this, we set the right-hand side of Equation 9 equal to the right-hand side of Equation 10, obtaining

$$-\frac{1}{v}(1 - (c + i)(1 - x))Y + \frac{1}{v}(C_0 + I_0 - (c + i)X^* + G^*) = \frac{k}{m}Y + \frac{1}{m}(L_0 - M^*)$$

Solving this equation for Y, we obtain

$$Y = \frac{1}{1 - (c + i)(1 - x) + \frac{vk}{m}}\left[C_0 + I_0 - \frac{v}{m}L_0 + \frac{v}{m}M^* - (c + i)X^* + G^*\right] \qquad (11)$$

By substituting Equation 11 into Equation 9 (or Equation 10) it is possible to solve for the equilibrium value of r also. However, since we shall have no occasion to use the resulting equation, we leave its derivation as an exercise for the student.

If we substitute a different set of values for C_0, I_0, L_0, M^*, X^*, and G^*, designated C_0', I_0', L_0', $M^{*\prime}$, $X^{*\prime}$, and $G^{*\prime}$, in Equation 11, we obtain a new value of Y designated by Y'; thus

$$Y' = \frac{1}{1 - (c + i)(1 - x) + \dfrac{vk}{m}} \left[C_0' + I_0' - \frac{v}{m} L_0' + \frac{v}{m} M^{*\prime} - (c + i)X^{*\prime} + G^{*\prime} \right] \quad (12)$$

Subtracting Equation 12 from Equation 11 and using the notation: $\Delta Y = Y - Y'$, $\Delta C_0 = C_0 - C_0'$, $\Delta I_0 = I_0 - I_0'$, $\Delta L_0 = L_0 - L_0'$, $\Delta M^* = M^* - M^{*\prime}$, $\Delta X^* = X^* - X^{*\prime}$, and $\Delta G^* = G^* - G^{*\prime}$, we obtain

$$\Delta Y = \frac{1}{1 - (c + i)(1 - x) + \dfrac{vk}{m}} \left[\Delta C_0 + \Delta I_0 - \frac{v}{m} \Delta L_0 + \frac{v}{m} \Delta M^* - (c + i)\Delta X^* + \Delta G^* \right] \quad (13)$$

From Equation 13, the following multipliers can be obtained (by setting all the increments in the brackets other than the relevant one equal to zero in each case):

1. *Multiplier for autonomous increase in the level of private investment:*

$$\frac{\Delta Y}{\Delta I_0} = \frac{1}{1 - (c + i)(1 - x) + \dfrac{vk}{m}} \quad (14)$$

2. *Multiplier for an increase in the level of government purchases:*

$$\frac{\Delta Y}{\Delta G^*} = \frac{1}{1 - (c + i)(1 - x) + \dfrac{vk}{m}} \quad (15)$$

3. *Multiplier for an increase in the level of net taxes:*

$$\frac{\Delta Y}{\Delta X^*} = \frac{-(c + i)}{1 - (c + i)(1 - x) + \dfrac{vk}{m}} \quad (16)$$

4. *Multiplier for an increase in the money stock:*

$$\frac{\Delta Y}{\Delta M^*} = \frac{\dfrac{v}{m}}{1 - (c + i)(1 - x) + \dfrac{vk}{m}} \quad (17)$$

Similar multipliers can be derived for the effect on income of autonomous shifts in the consumption function (ΔC_0) and in the demand for money function (ΔL_0).

The multipliers developed here differ in two respects from those developed in Chapter 6. To see the differences, let us examine the multiplier for a change in government purchases

$$\frac{\Delta Y}{\Delta G^*} = \frac{1}{1 - (c + i)(1 - x) + \dfrac{vk}{m}} \tag{15}$$

For purposes of comparison, the analogous multiplier from Chapter 6 (page 125) is

$$\frac{\Delta Y}{\Delta G^*} = \frac{1}{1 - c(1 - t)} \tag{18}$$

The first difference between the two multipliers is the presence of the additional term i in Equation 15. This is attributable to the appearance of the level of income in the investment Equation 2 of the present model. Thus $(c + i)(1 - x)$ is the marginal propensity to spend on both consumption and investment out of GNP in Equation 15, while $c(1 - t)$ is simply the marginal propensity to consume out of GNP in Equation 18.

The second and more important difference is reflected in the presence of the additional term vk/m in the denominator of Equation 15. It should be understood that Equation 15 measures the effect of an increase in the level of government purchases *with the money stock held constant.* If government purchases are increased, the resulting rise in income will increase the transactions demand for money. When the demand for money increases with a constant supply of money, the interest rate will rise. This rise in the interest rate will cause a reduction in investment, which will cancel out a portion of the increase in income caused by the initial rise in government spending, thereby reducing the size of the multiplier. Algebrically, it will be apparent that the term vk/m, which is positive, makes the denominator of Equation 15 larger, thereby cutting down the multiplier.

If investment, income, and the quantity of money demanded are denominated in billions of dollars and the interest rate is denominated as a percentage:

1. v represents the number of billions of dollars by which investment spending is increased as a result of a decline of one percentage point in the interest rate.
2. k represents the number of billions of dollars by which the transactions demand for money is increased as a result of an increase of $1 billion in GNP.

3. m represents the number of billions of dollars by which the quantity of money demanded is increased as a result of a decline of one percentage point in the interest rate.

Table 13–1 shows values of the multiplier for government purchases—or private investment, for which the multiplier (Equation 14) is the same as

TABLE 13–1

Values of Multipliers for Various Values of k, v, and m When $(c + i)(1 - x) = 0.4$

Values for $k = 0.1$

$$\frac{dY}{dG^*} = \frac{dY}{dI_0} = \frac{1}{1 - (c + i)(1 - x) + \frac{vk}{m}} \qquad \frac{dY}{dM^*} = \frac{\frac{v}{m}}{1 - (c + i)(1 - x) + \frac{vk}{m}}$$

Value of v	dY/dG^*: Value of m = 1	5	10	20	dY/dM^*: Value of m = 1	5	10	20
1..........	1.43	1.61	1.64	1.65	1.43	0.32	0.16	0.08
5..........	0.91	1.43	1.54	1.60	4.55	1.43	0.77	0.40
10..........	0.63	1.25	1.43	1.54	6.30	2.50	1.43	0.77
20..........	0.38	1.00	1.25	1.43	7.60	4.00	2.50	1.43

Values for $k = 0.25$

$$\frac{dY}{dG^*} = \frac{dY}{dI_0} = \frac{1}{1 - (c + i)(1 - x) + \frac{vk}{m}} \qquad \frac{dY}{dM^*} = \frac{\frac{v}{m}}{1 - (c + i)(1 - x) + \frac{vk}{m}}$$

Value of v	dY/dG^*: Value of m = 1	5	10	20	dY/dM^*: Value of m = 1	5	10	20
1..........	1.18	1.54	1.60	1.63	1.18	0.31	0.16	0.08
5..........	0.54	1.18	1.38	1.51	2.70	1.18	0.69	0.38
10..........	0.32	0.91	1.18	1.38	3.20	1.82	1.18	0.69
20..........	0.18	0.63	0.91	1.18	3.60	2.52	1.82	1.18

for government purchases—for various combinations of v, k, and m, on the assumption that the marginal propensity to spend out of GNP $[(c + i)(1 - x)]$ is equal to 0.4. It is apparent from the table that the multiplier becomes smaller as the size of v or k increases and becomes larger as the size of m increases. The reader will also see that these conclusions are apparent from the presence of the term vk/m in the denominator of Equation 15. The logic of these relationships can be made apparent in the following way:

1. The larger the value of k, the more the transactions demand for money will increase as a result of an increase in income caused by a rise in government purchases or an autonomous increase in private investment. The more the transaction demand for money increases, other things equal, the more will the interest rate rise and the larger will be the induced decline in investment expenditures.

2. The larger the value of m, the less will the interest rate have to rise to bring about the necessary economization of cash balances needed to meet the expanded transactions demand for money. The smaller the rise in the interest rate, other things equal, the less will be the induced decline in investment.

3. The larger the value of v, the greater will be the decline in investment expenditures resulting from a given rise in the interest rate.

It is useful to think of the term $1 - (c + i)(1 - x)$ in the denominator of Equation 15 as representing the *income effect* of a change in government purchases or an autonomous change in private investment and of the term vk/m as the *monetary feedback* which occurs when the stock money (M^*) is held constant. This monetary feedback, which reduces the size of the multiplier, was not included in the multiplier equations that were presented in Chapter 6 (such as Equation 18 reproduced above). The absence of the monetary feedback in the multipliers presented in Chapter 6 reflects an implicit assumption that when government purchases or private investment change, the central bank follows a policy of expanding or contracting the money supply sufficiently to keep the interest rate constant.[2] In other words, the multipliers of Chapter 6 may be interpreted as the multipliers that would apply if the central bank followed a policy of holding the interest rate constant, while the multipliers presented in this chapter would apply if the central bank followed a policy of holding the money supply constant.

The student will also recall that there was an extensive discussion of automatic fiscal stabilizers in Chapter 6. It should be apparent from the present discussion that there is also an *automatic monetary stabilizer* in the system which is represented by the term vk/m. To illustrate, let us suppose that there is an autonomous shift to the left of the marginal efficiency of investment schedule which causes the level of income to decline. The fall in income causes the transactions demand for money to decline. If we assume that the central bank takes no action—meaning, in the present context, that the stock of money remains unchanged—the decline in the demand for

[2] There are two alternative but less plausible assumptions that might account for the absence of a monetary feedback in the multipliers of Chapter 6. One is that the demand for money is infinitely elastic with respect to the interest rate so the additional transactions demand can be met without the necessity of any increase in the interest rate. This is the so-called "liquidity trap" situation depicted in Figure 11–2 and in Panel E of Figure 12–5, and it would be reflected in an infinitely large value of m, which would reduce the term vk/m to zero. The other possible assumption is that the marginal efficiency of investment schedule is completely inelastic with respect to the interest rate, so that a rise in the interest rate would have no effect on investment. This situation is depicted in Panel D of Figure 12–5 and would be reflected in zero value for v in the term vk/m.

money with a given supply of money will cause the interest rate to fall. As the interest rate declines, there will be some increase in investment partially offsetting the initial autonomous fall in investment. Thus, income will decline less than would have been the case had this automatic monetary adjustment not occurred. In the present model, the automatic fiscal stabilizers are represented by the tax rate, x, in the demoninator of the multiplier while the automatic monetary stabilizer is represented by the term vk/m.

Equation 17 is the multiplier applicable to an increase in the stock of money (ΔM^*). For convenience it is repeated here.

$$\frac{\Delta Y}{\Delta M^*} = \frac{\frac{v}{m}}{1 - (c + i)(1 - x) + \frac{vk}{m}} \tag{17}$$

It should be emphasized that this multiplier measures the effect of a *single once-and-for-all increase in the stock of money,* whereas the multiplier Equation 15 measures the effect of a *sustained* increase in the level of government purchases. Since comparing a change in the *stock* of money with a change in the *flow* of government spending is like comparing apples with oranges, it is not possible to draw any meaningful conclusions about the relative potency of fiscal and monetary policy by comparing the sizes of their respective multipliers.

The numerator (v/m) of Equation 17 may be thought of as reflecting the initial impact of monetary policy. An increase in the stock of money will initially cause the interest rate to decline, and this will stimulate investment spending, thus starting an expansion of income. The more the quantity of money demanded increases for a given decline in the interest rate—i.e., the larger the value of m—the less will a given increase in the stock of money cause the interest rate to fall and therefore the less powerful will monetary policy be. On the other hand, the larger is the value of v, the more will a given decline in the interest rate stimulate investment and the more powerful will monetary policy be. These relationships explain the fact that the parameters v and m appear in the numerator of Equation 17 in the form of the ratio v/m.

The denominator of the multiplier for monetary policy (Equation 17) is the same as that of the multiplier for government purchases (Equation 15). That is, the initial impact of monetary policy on investment, measured by the term v/k in the numerator of Equation 17, is subject to a multiplier identical with that applicable to an increase in government purchases. Thus, the multiplier for monetary policy, like that for fiscal policy, can be decomposed into an *income effect* measured by the term $1 - (c + i)(1 - x)$ and a *monetary feedback* measured by the term vk/m. The idea of a monetary feedback from monetary policy may seem rather confusing at first. However, the concept is quite easy to understand: monetary policy initially lowers interest rates and stimulates investment; as this causes a rise in income, however, the transactions demand for money increases, causing interest rates to rise back toward their original level.[3] The way in which this process works out will become more apparent later on in this chapter when we examine the rudimentary dynamics of monetary change.

Since the ratio v/m appears in both the numerator and denominator of Equation 17, it may not be immediately apparent whether an increase in this ratio will increase or decrease the multiplier for monetary policy. However, the weight of v/m in the denominator is muted by the presence of the term $1 - (c + i)(1 - x)$ and by the fact that v/m is multiplied by k, which is normally a fraction smaller than unity. Thus, v/m is more important in the numerator than in the denominator. As a consequence, an increase in v/m increases the size of the multiplier for monetary policy. Table 13–1 shows values of the multiplier (dY/dM^*) for changes in the stock of money for various values of v, m, and k on the assumption that the marginal propensity to spend out of GNP $[(c + i)(1 - x)]$ is equal to 0.4. It is apparent from this table that the multiplier increases as v increases and decreases as m increases. It will thus be noted that the effects of v and m on the multiplier for changes in the stock of money are precisely opposite to their effects on the multiplier for changes in government purchases (or for autonomous changes in private investment). On the other hand, an increase in k reduces the magnitude of both multipliers. It will be a useful exercise for the student to work out for himself the logic of these results.

It was pointed out above that the dimensions of monetary and fiscal action are not commensurable and that simple comparisons of their potency are generally meaningless. There are certain extreme situations, however, in which such comparisons may have validity—namely, situations in which the effect of one policy or the other reduces to zero. Such situations arise when extreme assumptions are made about the interest sensitivity of the demand for money. Two such situations may be considered.

1. The Demand for Money Is Infinitely Elastic with Respect to the Interest Rate. In this case, the value of m reaches infinity. By an examination of Equations 15 and 17, the student can see that the multipliers for a change in government purchases and a change in the stock of money reduce to

$$\frac{\Delta Y}{\Delta G^*} = \frac{1}{1 - (c + i)(1 - x)}$$

$$\frac{\Delta Y}{\Delta M^*} = 0$$

This is the "liquidity trap" case depicted in Figure 11–2 and Panel E of Figure 12–5. Here fiscal policy reaches its maximum strength for a given marginal propensity to spend out of GNP, since there is no monetary

[3] As we will see later in this chapter, it is conceivable that the feedback effect will be so powerful as to cause the interest rate to rise above the level from which it originally started.

feedback, while monetary policy becomes completely powerless because it does not change the interest rate. This is sometimes characterized as an *ultra-Keynesian* position, although in fact few persons of an avowedly Keynesian persuasion would accept it as even an approximate description of the world in which we live.

2. ***The Demand for Money Is Completely Inelastic with Respect to the Interest Rate.*** In this case, the value of m falls to zero, and the multiplier Equations 15 and 17 become[4]

$$\frac{\Delta Y}{\Delta G^*} = 0$$

$$\frac{\Delta Y}{\Delta M^*} = \frac{1}{k}$$

This is one interpretation of the extreme *quantity theory of money*—or what has, in some circles, come to be called the *monetarist* position. In this case, monetary policy is all-important and fiscal policy is completely powerless to affect income. Since a rise in the interest rate does not induce any economization of cash balances, an increase in government purchases (without a corresponding increase in taxes) must force up the interest rate enough to reduce private spending by an amount equal to the increase in government purchases, leaving total income unaffected.

Very few economists would adhere to either of these extreme positions. As indicated in Chapter 11, the overwhelming burden of evidence supports the view that the demand for money is sensitive to the interest rate but not infinitely sensitive. Thus, it is quite apparent that both fiscal and monetary policy possess power to affect income.

A NUMERICAL ILLUSTRATION

Many of the matters discussed above can usefully be clarified by means of a concrete numerical example. Suppose the model is as follows:

$$\begin{aligned}
C &= 0.7Y_d + 68 \\
I &= 0.1Y_d - 8r + 114 \\
X &= 0.25Y - 10 \\
Y_d &= Y - X \\
Y &= C + I + G^* \\
M_d &= 0.25Y - 10r + 5 \\
M_s &= M^* \\
M_d &= M_s \\
G^* &= 250 \\
M^* &= 205
\end{aligned}$$

[4] The indicated value of $\Delta Y/\Delta M^*$ is easily arrived at by multiplying the numerator and denominator of Equation 17 by m and then setting m equal to zero.

Substituting the indicated values of the parameters and exogenous variables ($c = 0.7, i = 0.1, x = 0.25, v = 8, k = 0.25, m = 10, C_0 = 68, I_0 = 114, L_0 = 5, M^* = 205, X^* = -10, G^* = 250$) in Equation 11 and solving, we obtain

$$Y = 1000$$

Starting with this value of Y, the values for taxes, disposable income, and consumption can be found in the following way

$$X = 0.25Y - 10 = 0.25(1000) - 10 = 240$$
$$Y_d = Y - X = 1000 - 240 = 760$$

Substituting M^* for M_d in the demand-for-money equation, we obtain

$$M^* = 0.25Y - 10r + 5$$

Substitution of the values for M^* and Y into this equation yields

$$205 = 0.25(1000) - 10r + 5$$

Solving for r, we obtain

$$r = 5\%$$

To obtain investment, substitute into the investment equation

$$I = 0.1Y_d - 8r + 114 = 0.1(760) - 8(5) + 114$$
$$I = 150$$

Saving and the government surplus are $B = X - G = -10$, and $S = Y_d - C = 160$.

All of these values are brought together in the first (original equilibrium) column of Table 13–2.

TABLE 13–2

Numerical Example of Multiplier for an Increase in Government Purchases (dollar amounts in billions)

	Original Equilibrium	*New Equilibrium**	*Change*
Gross National Product (Y)	$1,000	$1,020.0	+$20.0
Consumption (C)	600	610.5	+ 10.5
Investment (I)	150	147.5	− 2.5
Government purchases (G^*)	250	262.0	+ 12.0
Taxes (T)	240	245.0	+ 5.0
Disposable income (Y_d)	760	775.0	+ 15.0
Saving (S)	160	164.5	+ 4.5
Government surplus or deficit (−)($T-G$)	−10	−17.0	− 7.0
Money stock (M^*)	205	205.0	0.0
Interest rate (r)	5%	5.5%	+ 0.5%

* After an increase of $12 billion in the rate of government purchases.

Substituting the appropriate values of the parameters and exogenous variables into Equation 9 we obtain the following equation for the *IS curve for this model:*

$$r = -0.05Y + 55$$

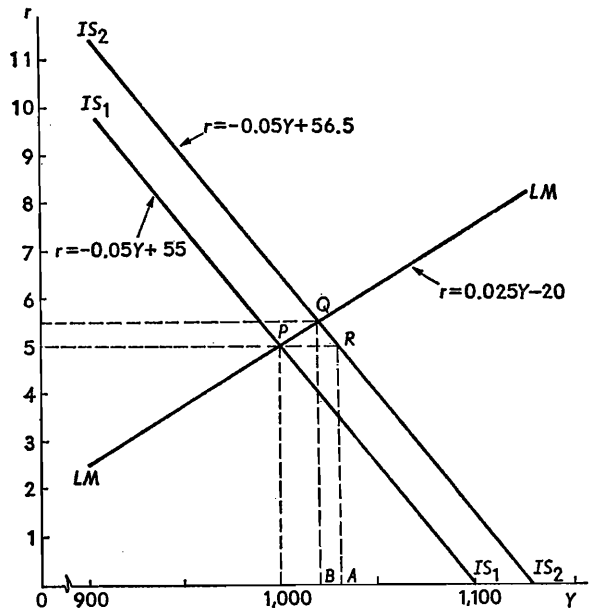
This equation is plotted as the line IS_1 in Figure 13–2. Similarly, substituting the appropriate values into Equation 10 we obtain the following equation for the *LM curve:*

$$r = 0.025Y - 20$$

This is plotted as line *LM* in Figure 13–2.

FIGURE 13–2

Effect of an Increase in Government Purchases



It will be noted that the lines IS_1 and *LM* intersect at the equilibrium values of $1,000 billion and 5 percent for *Y* and *r*. (Simultaneous algebraic solution of the equations for the *IS* and *LM* curves, of course, yields the same result.)

Effect of an Increase in Government Purchases

Starting from the equilibrium position calculated above, let us see what will be the effects of an increase of $12 billion a year in government purchases of goods and services. Substituting the assumed parameter values ($c = 0.7$, $i = 0.1$, $x = 0.25$, $v = 8$, $k = 0.25$, and $m = 10$) in Equation 15 we obtain the following value for the multiplier for government purchases

$$\frac{\Delta Y}{\Delta G^*} = 1\tfrac{2}{3} = 1.67$$

Applying this multiplier to the assumed $12 billion value for ΔG^*, we obtain

$$\Delta Y = \$20 \text{ billion}$$

Thus, the new equilibrium value for GNP (Y) is $1,020 billion. By following the procedure used above, the student can easily calculate the new values for all of the other variables. The results are shown in the second (new equilibrium) column of Table 13–2. The changes in all of the variables resulting from the $12 billion increase in the level of government purchases are shown in the last column of the table.

There are three points regarding these results that are especially worth noting:

1. The increase in government purchases causes the equilibrium interest rate to rise from 5 to 5.5 percent. This is because the increase in income raises the transactions demand for money while the supply of money is held constant at its original value of $205 billion.

2. The rate of investment is reduced by $2.5 billion, from $150 billion per year to $147.5 billion. There are two contradictory forces working on investment: the increase in disposable income tends to raise it, while the increase in the interest rate tends to lower it. However, the effect of the interest rate is the stronger of the two, causing investment to fall.

3. The increase of $12 billion in government purchases causes the government deficit to increase by $7 billion. The deficit increases by less than the increase in government purchases because the $20 billion increase in GNP causes tax collections to increase by $5 billion.

The increase in government purchases changes the *IS* curve. The equation of the new *IS* curve can be obtained by substituting the new value of G^* ($262 billion) along with the original values of the parameters and other exogenous variables in Equation 9. This yields the following equation for the *new IS curve:*

$$r = -0.05Y + 56.5$$

This curve is plotted as line IS_2 in Figure 13–2.

The *LM* curve is unaffected by the change in government purchases and therefore remains in its original position. The new equilibrium of the economy is depicted in Figure 13–2 by point *Q* where the *LM* curve (line

LM) and the new *IS* curve (line IS_2) intersect. The values of Y and r corresponding to point Q are, of course $1,020 billion and 5.5 percent.

It is useful to isolate the effect of the *monetary feedback* discussed earlier in this chapter. The monetary feedback is reflected in the term vk/m in the denominator of the multiplier (Equation 15). If this term is eliminated, the multiplier (analogous to the multipliers developed in Chapter 6) becomes

$$\frac{\Delta Y}{\Delta G^*} = \frac{1}{1 - (c + i)(1 - x)}$$

Evaluating this multiplier for $c = 0.7$, $i = 0.1$, and $x = 0.25$, we obtain

$$\frac{\Delta Y}{\Delta G^*} = 2.5$$

That is, the multiplier, in the absence of monetary feedback, is 2.5. The monetary feedback in this instance cuts the multiplier from 2.5 to 1.67, a reduction of 33.3 percent.

The monetary feedback is depicted graphically in Figure 13–2. The multiplier without monetary feedback is the increase in income that would be produced by an increase in government purchases if measures were taken to prevent the interest rate from rising. From Figure 13–2 it can be seen that if the *IS* curve shifted from IS_1 to IS_2 while the interest rate was held constant at its original value of 5 percent, the new equilibrium would be at point R, where income would be $1,030 billion and r would be 5 percent. Holding the money stock constant at $205 billion causes the interest rate to rise to 5.5 percent and income to increase only to $1,020 billion (at point Q). Thus, in this case the monetary feedback reduces the increase in GNP caused by a $12 billion increase in government purchases from $30 billion to $20 billion. The size of the monetary feedback (a negative $10 billion) is measured by the distance AB in Figure 13–2.

Effect of an Increase in the Money Stock

To find the effects of an increase in the money stock it is necessary to evaluate the multiplier (Equation 17). Substituting $c = 0.7$, $i = 0.1$, $x = 0.25$, $v = 8$, $k = 0.25$, and $m = 10$ into this expression, we obtain

$$\frac{\Delta Y}{\Delta M^*} = 1\tfrac{1}{3} = 1.33$$

Table 13–3 shows the effects on the economy of increasing the money stock by $12 billion. The "original equilibrium" of the economy as shown in the first column is the same as that in Table 13–2. In this case, however, government purchases are held at their original level of $250 billion and the money stock is increased from \$205 billion to \$217 billion. The increase in GNP is \$16 billion (1.33 × \$12 billion). The values of the variables in the new equilibrium are shown in the second column of Table 13–3. The student should calculate these values for himself, using the procedure outlined earlier in this chapter. The changes in the variables caused by the increase in the money stock are shown in the last column of Table 13–3. Three points are specifically worth noting.

1. The interest rate declines from 5 to 4.2 percent. The increase in the money stock tends to lower the interest rate, while the resulting rise in GNP, by increasing the demand for money, tends to raise it. The first effect,

TABLE 13–3

Numerical Example of the Multiplier for an Increase in the Stock of Money (dollar amounts in billions)

	Original Equilibrium	*New Equilibrium**	*Change*
Gross national product (Y)	\$1,000	\$1,016.0	+\$16.0
Consumption (C)	600	608.4	+ 8.4
Investment (I)	150	157.6	+ 7.6
Government purchases (G^*)	250	250.0	0.0
Taxes (T)	240	244.0	+ 4.0
Disposable income (Y_d)	760	772.0	+ 12.0
Saving (S)	160	163.6	+ 3.6
Government surplus or deficit (−)($T-G$)	−10	−6.0	+ 4.0
Money stock (M^*)	205	217.0	+ 12.0
Interest rate (r)	5%	4.2%	− 0.8%

* After an increase of \$12 billion in the money stock.

however, predominates, with the result that the interest rate falls. (As will be shown later in this chapter, this need not always be the case.)

2. Investment spending increases by \$7.6 billion, from a rate of \$150 billion per year to \$157.6 billion. In this case investment is doubly stimulated by a fall in the interest rate and a rise in disposable income.

3. The government deficit declines from \$10 billion to \$6 billion. This is because government purchases are unchanged while tax collections increase by \$4 billion as a result of the rise in GNP.

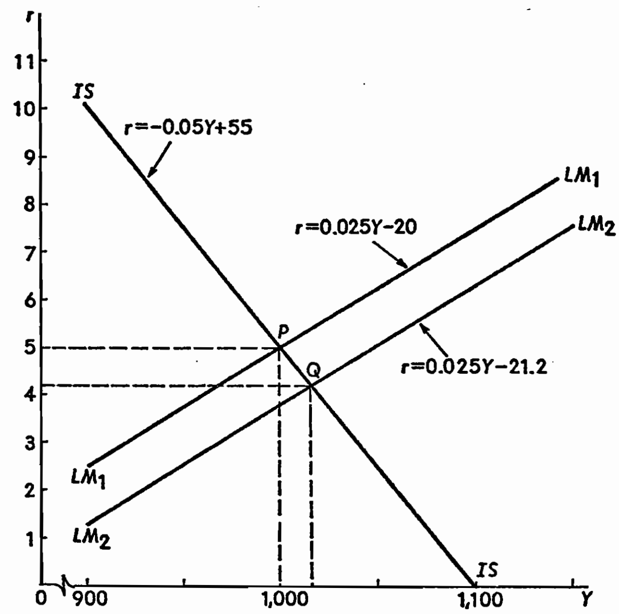
The effect of the increase of \$12 billion in the money stock can be shown by means of *IS* and *LM* curves. The *IS* curve is unaffected by the change in money stock; it is depicted by the line *IS* in Figure 13–3—which is, of course, exactly the same as the line IS_1 in Figure 13–2. Its equation, repeated here for convenience, is

$$r = -0.05Y + 55$$

The *LM* curve, however, is shifted by the change in the money stock. The equation of the new *LM* curve is obtained by substituting the new value of M^* (\$217 billion) along with the original values of the parameters and

FIGURE 13–3

Effect of an Increase in the Stock of Money



other exogenous variables in Equation 10. This yields the following equation for the *new LM curve:*

$$r = 0.025Y - 21.2$$

The original *LM* curve is represented by line LM_1 in Figure 13–3, and the new *LM* curve is represented by line LM_2. The equilibrium after the increase of \$12 billion is at point *Q* where the lines *IS* and LM_2 intersect. The values of *Y* and *r* at this point are, of course, \$1,016 billion and 4.2 percent.

DYNAMIC ANALYSIS

It will be useful to examine a relatively simple dynamic version of the model we have been studying. Such a version can be represented by the following set of equations.

$$C_t = cY_t^d + C_0 \quad (19)$$
$$I_t = iY_t^d - vr_t + I_0 \quad (20)$$
$$X_t = xY_t + X^* \quad (21)$$
$$Y_t^d = Y_t - X_t \quad (22)$$
$$Y_t = C_{t-1} + I_{t-1} + G_{t-1}^* \quad (23)$$
$$M_t^d = kY_t - mr_t + L_0 \quad (24)$$
$$M_t^s = M_t^* \quad (25)$$
$$M_t^d = M_t^s \quad (26)$$

The only time lag in this model is the output lag—the lag in the adjustment of production to sales—which is reflected in Equation 23. This is admittedly a tremendous oversimplification; in reality there would be complex lags in many of the other equations. Moreover, even this dynamic version of the *IS–LM* model does not contain the investment accelerator discussed in Chapter 9. Nevertheless, in spite of its simplicity, some valuable insights can be derived from an examination of this model.

As a first step in analyzing the model, it is useful to shift each of the Equations 19, 20, 21, and 22 back one period. This process yields

$$C_{t-1} = cY_{t-1}{}^{d} + C_0 \qquad (19')$$
$$I_{t-1} = iY_{t-1}{}^{d} - vr_{t-1} + I_0 \qquad (20')$$
$$X_{t-1} = xY_{t-1} + X^* \qquad (21')$$
$$Y_{t-1}{}^{d} = Y_{t-1} - X_{t-1} \qquad (22')$$

Equations 19′, 20′, 21′, and 22′ can now be substituted into Equation 23. Upon doing this and collecting terms, and assuming $G_t{}^*$ is a constant designated simply as G^*, we obtain

$$Y_t = (c + i)(1 - x)Y_{t-1} - vr_{t-1} + C_0 + I_0 - (c + i)X^* + G^* \qquad (27)$$

Substituting from Equations 25 and 26 into Equation 24 and assuming $M_t{}^*$ is a constant designated as M^*, we obtain

$$M^* = kY_t - mr_t + L_0$$

Shifting this equation back one period in time and solving for r_{t-1}, we have

$$r_{t-1} = \frac{k}{m}Y_{t-1} + \frac{1}{m}(L_0 - M^*) \qquad (28)$$

We can eliminate r_{t-1} from Equation 27 by substituting for it the value given by Equation 28; upon substituting and collecting terms this yields

$$Y_t = \left[(c + i)(1 - x) - \frac{vk}{m}\right] Y_{t-1} + C_0 + I_0 - \frac{v}{m}(L_0 - M^*) - (c + i)X^* + G^* \qquad (29)$$

This is a difference equation in Y—it shows the value of Y in period t as a function of its value in period $t - 1$ and of the parameters and exogenous variables of the model. It is apparent that if the value of Y for a single period is known and if the parameters and exogenous variables are given and remain constant, values of Y for all ensuing periods can be obtained by using Equation 29.

An equilibrium position for this model means that Y remains constant from period to period. We can obtain the equilibrium value ($\bar{Y}$) of Y by assuming that $Y_t = Y_{t-1} = \bar{Y}$ in Equation 29. This yields

$$\bar{Y} = \left[(c + i)(1 - x) - \frac{vk}{m}\right]\bar{Y} + C_0 + I_0 - \frac{v}{m}(L_0 - M^*) - (c + i)X^* + G^* \quad (30)$$

Solving this equation for the equilibrium value of Y, we obtain

$$\bar{Y} = \frac{1}{1 - (c + i)(1 - x) + \frac{vk}{m}}\left[C_0 + I_0 - \frac{v}{m}L_0 + \frac{v}{m}M^* - (c + i)X^* + G^*\right] \quad (31)$$

The student will note that this is exactly the same as Equation 11, which gives the equilibrium value of Y for the static model presented earlier in this chapter.

Numerical Illustrations

It will be useful to work out in a dynamic context the numerical examples presented earlier in the chapter, showing the effects of a change in government purchases and in the stock of money. We shall use the same values of the parameters that were assumed earlier: $c = 0.7$, $i = 0.1$, $x = 0.25$, $v = 8$, $k = 0.25$, $m = 10$, $C_0 = 68$, $T^* = -10$, $I_0 = 114$, and $L_0 = 5$. Initially, it is assumed that $G^* = 250$ and $M^* = 205$. The values of the endogenous variables in the initial equilibrium, corresponding to these values of G^* and M^* are shown in the Original Equilibrium columns of Tables 13–4 and 13–5. The entries in these columns are, of course, the same as the entries in the Original Equilibrium columns of Tables 13–2 and 13–3.

In Table 13–4, it is assumed that the level of government purchases (G^*) rises by \$12 billion to \$262 billion in period one and remains there. The table shows the dynamic process by which the endogenous variables move toward their new equilibrium values, which are shown in the last column of the table. These values are, of course, the same as those shown in the New Equilibrium column of Table 13–2.

Table 13–5 shows the process by which the variables move from one equilibrium position toward another when the money stock (M^*) is increased. It is assumed in this illustration that government purchases (G^*) remain at their original level of \$250 billion, while the money stock is increased by \$12 billion to \$217 billion in period one and remains at this higher level. The entries in the New Equilibrium column of Table 13–5 are, of course, the same as the entries in the corresponding column of Table 13–3.

The behavior of the interest rate is especially worth noting in Table 13–5. The interest rate falls sharply from 5 to 3.8 percent in period one,

TABLE 13-4

Dynamics of an Increase of $12 Billion in Government Purchases

	How Computed	Original Equilibrium	Time Period (t) 1	2	3	4	New Equilibrium
Y_t	$C_{t-1} + I_{t-1} + G_{t-1}^*$	$1,000	$1,000	$1012.0	$1016.80	$1018.720	$1020.0
C_t	$0.7Y_t^d + 68$	600	600	606.3	608.82	609.828	610.5
I_t	$0.1Y_t^d - 8r_t + 114$	150	150	148.5	147.90	147.660	147.5
G_t^*	Autonomous	250	262	262.0	262.00	262.000	262.0
X_t	$0.25Y_t - 10$	240	240	243.0	244.20	244.680	245.0
Y_t^d	$Y_t - X_t$	760	760	769.0	772.60	774.040	775.0
S_t	$Y_t^d - C_t$	160	160	162.7	163.78	164.212	164.5
$X_t - G_t^*$	$X_t - G_t^*$	−10	−22	−19.0	−17.80	−17.320	−17.0
M_t^*	Autonomous	205	205	205.0	205.00	205.000	205.0
r_t	$\frac{1}{10}[0.25Y_t + 5 - M_t^*]$	5%	5%	5.3%	5.42%	5.468%	5.5%

TABLE 13-5

Dynamics of an Increase of $12 Billion in the Money Stock

	How Computed	Original Equilibrium	Time Period (t) 1	2	3	4	New Equilibrium
Y_t	$C_{t-1} + I_{t-1} + G_{t-1}^*$	$1000	$1000.0	$1009.60	$1013.440	$1014.9760	$1016.0
C_t	$0.7Y_t^d + 68$	600	600.0	605.04	607.056	607.8624	608.4
I_t	$0.1Y_t^d - 8r_t + 114$	150	159.6	158.40	157.920	157.7280	157.6
G_t^*	Autonomous	250	250.0	250.00	250.000	250.0000	250.0
X_t	$0.25Y_t - 10$	240	240.0	242.40	243.360	243.7440	244.0
Y_t^d	$Y_t - X_t$	760	760.0	767.20	770.080	771.2320	772.0
S_t	$Y_t^d - C_t$	160	160.0	162.16	163.024	163.3696	163.6
$X_t - G_t^*$	$X_t - G_t^*$	−10	−10.0	−7.60	−6.640	−6.2560	−6.0
M_t^*	Autonomous	205	217.0	217.00	217.000	217.0000	217.0
r_t	$\frac{1}{10}[0.25Y_t + 5 - M_t^*]$	5%	3.8%	4.04%	4.136%	4.1744%	4.2%

when the rise in the money stock occurs. However, in succeeding periods, as the lower interest rate stimulates investment, causing income and the transactions demand for money to increase, the interest begins to rise toward its new equilibrium value of 4.2 percent.

Stability Conditions

Let us define the variable y_t to be the deviation of GNP in period t from its equilibrium value; that is,

$$y_t = Y_t - \bar{Y}$$

The same definition holds, of course, for other periods:

$$y_{t-1} = Y_{t-1} - \bar{Y},\ y_{t-2} = Y_{t-2} - \bar{Y}$$

and so on.

If we subtract Equation 30 from Equation 29, we obtain

$$Y_t - \bar{Y} = \left[(c + i)(1 - x) - \frac{vk}{m}\right](Y_{t-1} - \bar{Y})$$

or

$$y_t = \left[(c + i)(1 - x) - \frac{vk}{m}\right] y_{t-1}$$

If we are given the value, Y_0, of GNP in an initial period, we can define $y_0 = Y_0 - \bar{Y}$ ($\bar{Y}$, of course, being a number whose value can be calculated from Equation 30 for given values of the parameters and exogenous variables). Then we have

$$y_1 = \left[(c + i)(1 - x) - \frac{vk}{m}\right] y_0$$

$$y_2 = \left[(c + i)(1 - x) - \frac{vk}{m}\right] y_1 = \left[(c + i)(1 - x) - \frac{vk}{m}\right]^2 y_0$$

Or, in general,

$$y_t = \left[(c + i)(1 - x) - \frac{vk}{m}\right]^t y_0$$

Since $y_t = Y_t - \bar{Y}$ and $y_0 = Y_0 - \bar{Y}$, this equation can be written

$$Y_t - \bar{Y} = \left[(c + i)(1 - x) - \frac{vk}{m}\right]^t (Y_0 - \bar{Y})$$

or

$$Y_t = \left[(c + i)(1 - x) - \frac{vk}{m}\right]^t (Y_0 - \bar{Y}) + \bar{Y} \tag{32}$$

where

$$\bar{Y} = \frac{1}{1 - (c + i)(1 - x) + \dfrac{vk}{m}} \left[C_0 + I_0 - \frac{v}{m} L_0 + \frac{v}{m} M^* - (c + i)X^* + G^*\right] \tag{31}$$

Thus, if the values of all of the parameters and exogenous variables are given and if Y starts from a given arbitrary value, Y_0, in period zero, the values of Y in all successive periods, $t = 1$, $t = 2$, etc., can be calculated from Equations 32 and 31. The system will be stable if, starting from an arbitrary value, Y_0, it approaches its equilibrium value, $\bar{Y}$, with the passage of time. This will be the case if

$$(c + i)(1 - x) - \frac{vk}{m} < 1$$

since, if this condition is fulfilled, the first term of Equation 32 will approach zero as t approaches infinity, leading to the result

$$[Y_t]_{t \to \infty} = \bar{Y}$$

Thus, the stability condition for the system is

$$(c + i)(1 - x) - \frac{vk}{m} < 1 \tag{33}$$

Since the multiplier analysis assumes that the economy starts from a position of equilibrium, we can use the ideas developed above to work out a dynamic version of the multiplier by assuming that the initial value, Y_0, from which the economy starts, is an equilibrium value. For example, let us suppose that the economy is in equilibrium initially at $\bar{Y}_0$ with government purchases of $G_0{}^*$. That is

$$\bar{Y}_0 = \frac{1}{1 - (c + i)(1 - x) + \dfrac{vk}{m}} \left[C_0 + I_0 - \frac{v}{m} L_0 + \frac{v}{m} M^* - (c + i)X^* + G_0{}^*\right] \tag{34}$$

Now suppose government purchases increase to $G_1{}^*$. The equilibrium, $\bar{Y}_1$, will be

$$\bar{Y}_1 = \frac{1}{1 - (c + i)(1 - x) + \frac{vk}{m}}\left[C_0 + I_0 - \frac{v}{m}L_0 + \frac{v}{m}M^* - (c + i)X^* + G_1^*\right] \quad (35)$$

Equation 32 becomes

$$Y_t = \left[(c + i)(1 - x) - \frac{vk}{m}\right]^t(\bar{Y}_0 - \bar{Y}_1) + \bar{Y}_1$$

Subtracting $\bar{Y}_0$ from both sides of this equation, we have

$$Y_t - \bar{Y}_0 = -\left[(c + i)(1 - x) - \frac{vk}{m}\right]^t(\bar{Y}_1 - \bar{Y}_0) + \bar{Y}_1 - \bar{Y}_0 \quad (36)$$

Subtracting Equation 34 from Equation 35, we obtain

$$\bar{Y}_1 - \bar{Y}_0 = \frac{G_1^* - G_0^*}{1 - (c + i)(1 - x) + \frac{vk}{m}} \quad (37)$$

Substituting Equation 37 into Equation 36 and simplifying, we obtain

$$Y_t - \bar{Y}_0 = \frac{1 - \left[(c + i)(1 - x) - \frac{vk}{m}\right]^t}{1 - (c + i)(1 - x) + \frac{vk}{m}}(G_1^* - G_0^*)$$

Letting $\Delta Y_t = Y_t - \bar{Y}_0$ and $\Delta G^* = G_1^* - G_0^*$, we have

$$\Delta Y_t = \frac{1 - \left[(c + i)(1 - x) - \frac{vk}{m}\right]^t}{1 - (c + i)(1 - x) + \frac{vk}{m}}\Delta G^*$$

If we designate by ΔY the value of ΔY_t as t becomes infinitely large and if the stability condition $(c + i)(1 - x) - vk/m < 1$ is satisfied, we have

$$\Delta Y = \frac{1}{1 - (c + i)(1 - x) + \frac{vk}{m}}\Delta G^*$$

This is the same as the static multiplier for a change in government purchases (Equation 15) developed earlier in this chapter.

It will be recalled that the stability condition for a multiplier model without monetary feedback, as presented in Chapter 6, was that the marginal propensity to spend out of GNP could not exceed unity. In the present context, this condition would be

$$(c + i)(1 - x) < 1$$

On the other hand, the stability condition (Equation 33) for the present model can be written as

$$(c + i)(1 - x) < 1 + \frac{vk}{m}$$

Since the monetary feedback exerts a stabilizing influence, the model with monetary feedback will be stable even if the marginal propensity to spend out of GNP is greater than unity, provided it is less than unity plus the term vk/m, which measures the strength of the monetary feedback.

The stability condition can also be related to the slopes of the *IS* and *LM* curves. From Equations 9 and 10, the slope of the *IS* curve is

$$\frac{\Delta r}{\Delta Y} = -\frac{1}{v}(1 - (c + i)(1 - x))$$

and the slope of the *LM* curve is

$$\frac{\Delta r}{\Delta Y} = \frac{k}{m}$$

The stability condition (Equation 33) can be rewritten as follows

$$-\frac{vk}{m} < 1 - (c + i)(1 - x)$$

Multiplying both sides of this inequality by -1 and reversing the inequality signs as this multiplication requires, we have

$$\frac{vk}{m} > -(1 - (c + i)(1 - x))$$

Dividing both sides of this inequality by v, we obtain

$$\frac{k}{m} > -\frac{1}{v}(1 - (c + i)(1 - x))$$

Since the left-hand side of this inequality is the slope of the *LM* curve and the right-hand side is the slope of the *IS* curve, the stability condition can be restated as

$$\text{Slope of } LM \text{ curve} > \text{Slope of } IS \text{ curve}$$

In our earlier numerical illustrations, the relevant parameters had the following values: $c = 0.7$, $i = 0.1$, $x = 0.25$, $v = 8$, $k = 0.25$, $m = 10$. In this case the stability condition (Equation 33) was satisfied, since

$$(c + i)(1 - x) - \frac{vk}{m} = 0.4$$

It is also apparent from Figures 13–2 and 13–3 that the slope of the *IS* curve is less than the slope of the *LM* curve, since the slope of the *IS* curve is negative while that of the *LM* curve is positive.

As was pointed out earlier in this chapter, the slope of the *LM* curve is *always* positive. However, it is possible to have a stable economy in which the *IS* curve is positively sloped, provided its slope is less than that of the *LM* curve. It will be instructive to consider a numerical example in which this is the case. Consider the following set of equations

$$\begin{aligned}
C_t &= 0.9Y_t^d \\
I_t &= 0.7Y_t^d - 20r_t - 282 \\
X_t &= 0.25Y_t + 30 \\
Y_t^d &= Y_t - X_t \\
Y_t &= C_{t-1} + I_{t-1} + G^* \\
M_t^d &= 0.30Y_t - 10r_t - 30 \\
M_t^s &= M^* \\
M_t^d &= M_t^s \\
G^* &= 270 \\
M^* &= 200
\end{aligned}$$

The values of the parameters and exogenous variables in this case are: $c = 0.9$, $i = 0.7$, $x = 0.25$, $v = 20$, $k = 0.30$, $m = 10$, $C_0 = 0$, $I_0 = -282$, $X^* = 30$, $L_0 = -30$, $M^* = 200$, $G^* = 270$. To find the equilibrium value of Y, we substitute these values into Equation 11, obtaining

$$Y = \$1{,}000 \text{ million}$$

The equilibrium values of all of the endogenous variables are shown in the Original Equilibrium column of Table 13–6.

The equations of the *IS* and *LM* curves, obtained by substituting the relevant values into Equations 9 and 10, are

$$r = 0.01Y - 3 \qquad IS \text{ curve}$$

$$r = 0.03Y - 23 \qquad LM \text{ curve}$$

The stability condition is satisfied, since

$$(c + i)(1 - x) - \frac{vk}{m} = 0.6$$

or, alternatively,

$$\text{Slope of } LM \text{ curve} = 0.03 > \text{Slope of } IS \text{ curve} = 0.01$$

The *IS* curve is line *IS* while the *LM* curve is line LM_1 in Figure 13–4. These two lines intersect at point *P*. The equilibrium values of *Y* and *r* corresponding to this point are $Y = \$1{,}000$ billion and $r = 7$ percent.

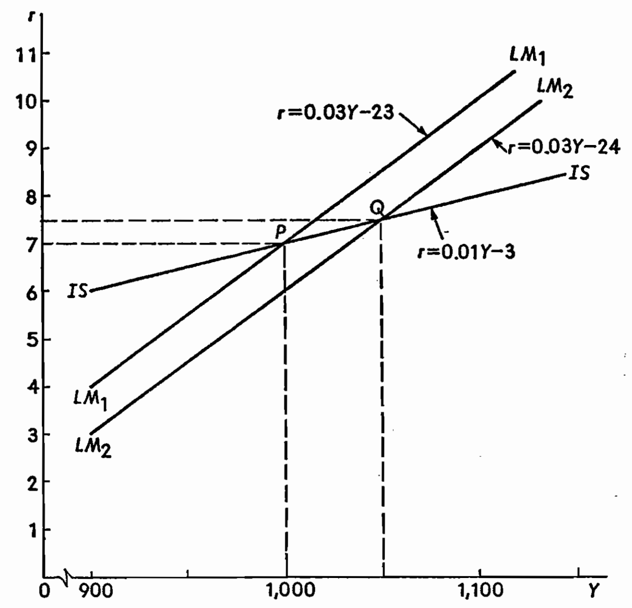
TABLE 13-6

Dynamics of an Increase of $10 Billion in the Money Stock with an Upward Sloping *IS* Curve

	How Computed	Original Equilibrium	Time Period (t) 1	2	3	4	New Equilibrium
Y_t	$C_{t-1} + I_{t-1} + G_{t-1}^*$	$1000	$1000	$1020.0	$1032.0	$1039.20	$1050.00
C_t	$0.9Y_t^d$	648	648	661.5	669.6	674.46	681.75
I_t	$0.7Y_t^d - 20r_t - 282$	82	102	100.5	99.6	99.06	98.25
G_t^*	Autonomous	270	270	270.0	270.0	270.00	270.00
X_t	$0.25Y_t^d + 30$	280	280	285.0	288.0	289.80	292.50
Y_t^d	$Y_t - X_t$	720	720	735.0	744.0	749.40	757.50
S_t	$Y_t^d - C_t$	72	72	73.5	74.4	74.94	75.75
$X_t - G_t^*$	$X_t - G_t^*$	10	10	15.0	18.0	19.80	22.50
M_t^*	Autonomous	200	210	210.0	210.0	210.00	210.00
r_t	$\frac{1}{10}[0.30Y_t - 30 - M_t^*]$	7%	6%	6.6%	6.96%	7.176%	7.5%

FIGURE 13-4

Effect of an Increase in the Stock of Money with an Upward Sloping *IS* Curve



It will be instructive to examine the behavior of the economy when the money stock is increased in this case. The multipliers for a change in government purchases (ΔG^*) and for a change in the stock of money (ΔM^*), obtained by substituting the appropriate values in Equations 15 and 17, are

$$\frac{\Delta Y}{\Delta G^*} = 2.5$$

$$\frac{\Delta Y}{\Delta M^*} = 5$$

In Table 13–6 it is assumed that the money stock is increased by $10 billion, from $200 billion to $210 billion, in period one. The table shows the dynamics of the expansion caused by this increase, and, in the last column, the new equilibrium that will ultimately be reached.

The increase of $10 billion in the money stock shifts the *LM* curve from LM_1 to LM_2 in Figure 13–4. The equation of the new *LM* curve is

$$r = 0.03Y - 24$$

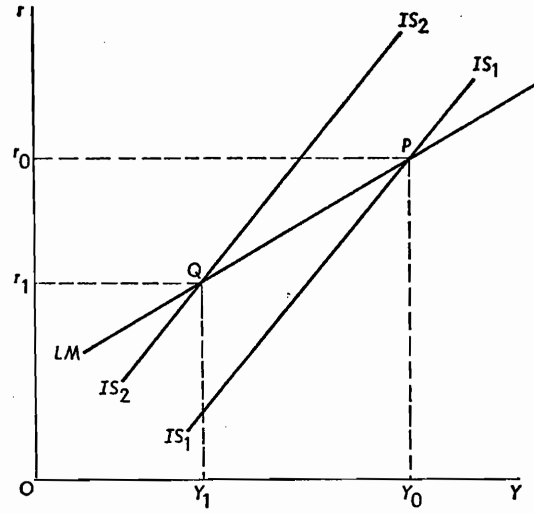
The new equilibrium is at point *Q* where the line LM_2 intersects the line *IS*. The new equilibrium values corresponding to this point are $Y = \$1{,}050$ billion and $r = 7.5$ percent.

The distinctive feature of this illustration is that an increase in the stock of money causes the equilibrium interest rate to rise. This is clearly apparent in Table 13–6. In period one the interest rate falls sharply from its initial equilibrium value of 7 percent, to 6 percent. In succeeding periods, the expansion of income increases the transactions demand for money. By period four, the interest rate has risen above its original level of 7 percent, and its new equilibrium value, shown in the last column, is 7.5 percent.

For purposes of clarification, an unstable situation is depicted in Figure 13–5. In this case the slope of the *IS* curve is positive and greater than that

FIGURE 13–5

Illustration of an Unstable Economy



of the *LM* curve. It is possible to see in common sense terms why this economy is unstable. Suppose initially the *IS* curve is the line IS_1 while the *LM* curve is the line *LM*. The economy is in equilibrium at point *P* where these lines intersect, with $Y = Y_0$ and $r = r_0$. Now an increase in government purchases occurs, raising the *IS* curve from IS_1 to IS_2.[5] The new equilibrium point will be at *Q*, where the new *IS* curve intersects the *LM* curve, with equilibrium values $Y = Y_1$ and $r = r_1$. The instability of the system is apparent from the fact that the equilibrium value of *Y* with the higher level of government purchases is *below* the original equilibrium value. It is quite clear that the increase in government purchases will cause income to increase, thereby moving it *away from* the new equilibrium value rather than toward that value.

[5] It is apparent from Equation 9 that an increase in government purchases (G^*) will cause a vertical shift in the *IS* curve by increasing the size of its intercept on the *r*-axis.

Chapter 14

SOME REFINEMENTS OF AGGREGATE DEMAND THEORY

In several respects, the analysis presented in previous chapters is considerably oversimplified, particularly in its treatment of the financial sector of the economy. The purpose of this chapter is to round out the analysis by taking account of some of the complications that have been left out up to this point.

THE STRUCTURE OF INTEREST RATES

In our earlier discussion, we have considered the role of, and the determination of, "the" interest rate, thereby implicitly assuming the existence of only one interest rate in the economy. This is obviously a gross oversimplification; in fact there are hundreds of interest rates. These interest rates are not the same, although they are related. In this section, we will discuss some of the relationships that exist among interest rates.

Debt contracts differ among themselves in a number of respects. Some of the main dimensions that can be used in classifying them are as follows:

1. *Maturity.* The time that elapses between the time a debt contract is entered into and the time when the borrower agrees to repay the debt varies from a day or two to a period of many years.[1] Usually, the maturity of a debt that is meaningful is the period that remains until maturity at the time of observation, rather than the time that was scheduled to elapse between the original issuance of the contract and its ultimate repayment. Thus, a bond issued in 1950 and due for repayment 30 years later in 1980 had a maturity of 12 years in 1968. Some debts, such as demand deposits in a bank (which are debts of the bank to the depositors) are due on demand—that is, whenever the holder wants to cash them in—and do not therefore have a specific maturity.

2. *Marketability.* Many debts, such as bonds issued by governmental units or private corporations, have organized resale markets on which they

[1] The extreme case of long maturity is a *consol,* which is a debt contract under which the borrower agrees to pay interest indefinitely into the future without ever repaying the principal of the debt. Thus, the maturity of a consol is infinitely long. We have had several occasions to use consols for illustrative purposes in earlier chapters.

can be sold by their owners, prior to maturity, at prices that vary in such a way as to keep their yields in line with prevailing interest rates. Such debts may be described as *marketable.* Other debt contracts, including most loans made by banks and other financial institutions, cannot be sold by the lenders prior to maturity and are therefore termed *nonmarketable.*[2] Some debts, while not marketable, are *redeemable* on demand at a specified price by the holder. For example, bank deposits can be redeemed at face value in cash at the bank, and U.S. Savings Bonds can be redeemed at face value plus accrued interest at the U.S. Treasury.

3. *Risk.* Debt contracts differ with respect to risk of default—that is, risk that the borrower will be unable to pay the interest when due or repay the principal at maturity. Debt issues of the U.S. Treasury are free of default risk, since the federal government has the power to issue money to pay off its debt if necessary. Private debt issued by individuals and business enterprises is subject to a degree of default risk that varies according to the financial and economic status of the borrower. Since state and local government units do not possess money-creating powers, their debt is also subject to a degree of risk that varies with the financial condition of the unit, the adequacy of its tax base, and other factors. Some types of debt are guaranteed or insured by the federal government or its agencies; this reduces the degree of default risk attaching to such debt. For example, the Federal Housing Agency operates a program of mortgage insurance designed to reduce the risk attaching to mortgages that come under the program.

4. *Tax status.* The main provision of the tax laws that affects the status of debt instruments is the provision which exempts state and local government debt from the federal income tax.

5. *Legal and administrative restrictions.* There are a variety of restrictions which limit the amount of interest that can be paid on debt obligations. Many states have usury laws which set upper limits on the rate of interest that can be charged on various kinds of debt. There are often statutory or administrative limitations on the interest that can be paid on state and local government security issues. The Federal Reserve and other federal agencies empowered to regulate banks and other financial institutions have authority to establish ceilings applicable to the interest rates that can be paid on various classifications of time and savings deposits of banks and savings and loan associations. The federal housing agencies are empowered to limit interest rates that can be charged on mortgage loans guaranteed or insured by the federal government. There is a legal limit of 4¼ percent on the interest rate the federal government can pay on new issues of its securities having a maturity longer than seven years.

All of the features listed above, some of which are inherent in the nature of debt contracts and some of which reflect the working of legal provisions

---

[2] Loans made by banks that are members of the Federal Reserve may under certain conditions be discounted at (i.e., sold to) the Federal Reserve or pledged as collateral for advances from the Federal Reserve.

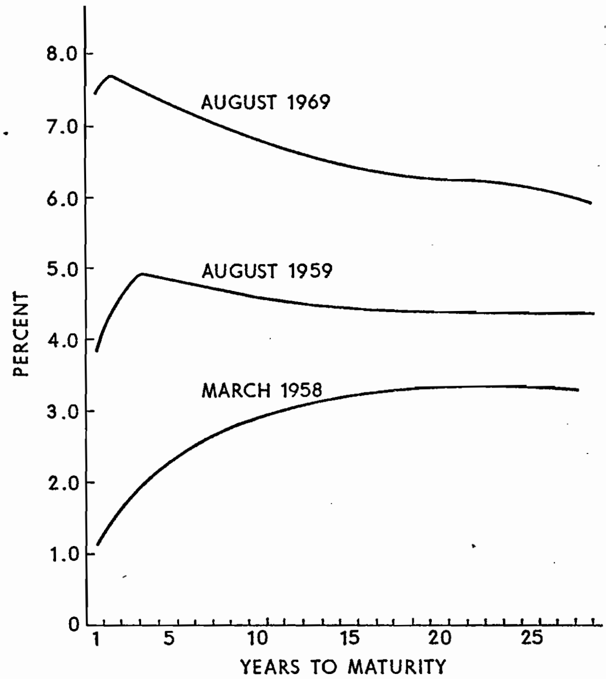
and constraints, influence the structure of interest rates on various types of debt.

The Maturity Structure of Interest Rates

Interest rates vary in a reasonably systematic way according to the maturity of debt instruments. In order to isolate the effects of debt maturity, it is necessary to abstract from the multitude of other factors affecting interest rates. The usual way of accomplishing this is to confine the analysis

FIGURE 14–1

Maturity Structure of Interest Rates
(March 1958, August 1959, and August 1969)



to interest rates on marketable U.S. government debt. As indicated above, such debt is completely free of default risk, since the federal government can, if necessary, issue money to pay off its debt.

Figure 14–1 presents so-called yield curves for U.S. government securities as compiled by the Treasury Department for the months of March 1958, August 1959, and August 1969.[3] Each of these yield curves shows how the

[3] These yield curves are taken from issues of the *Treasury Bulletin* for April 1958, September 1959, and September 1969, respectively.

yields to maturity on U.S. government securities varied with the maturity of the securities on the indicated date.[4]

The Expectations Hypothesis

The most common explanation of the maturity structure of interest rates is based on the so-called *expectations hypothesis*.[5] To illustrate this hypothesis, we make the following assumptions:

1. All investors have definite expectations with respect to future short-term interest rates, and these expectations are held with complete confidence.
2. The objective of investors is to maximize expected profits, and they are prepared to transfer funds freely from one maturity to another in order to achieve this objective.
3. There are no costs associated with investment and disinvestment in securities.

Suppose we take the current short-term (one-period) interest rate, R_1, to be given and assume that investors have definite expectations concerning the sequence of one-period rates, $r_2, r_3, \ldots, r_n$, which will prevail in coming periods.[6] Then if $1 is invested in a one-period security and the proceeds

---

[4] The yields to maturity are calculated by use of equation (2) of Chapter 11. This equation is repeated here for convenience.

$$V = cP\frac{1-(1+r)^{-n}}{r} + P(1+r)^{-n}$$

where V is the price of the security, P is its face (or par) value, c is the coupon interest rate, n is the number of years to maturity, and r is the yield to maturity. With the face value, coupon rate, and period to maturity given by the terms of the security, and with the current price given by the market, the above equation can be used to find the yield to maturity. For example, if a security has a face value of $1,000, a coupon rate of 5 percent, and 16 years remaining to maturity, and if it is currently selling at a price of $895.94, its yield to maturity is 6 percent, as calculated from the above equation. In practice, most bonds pay interest twice a year; when this is the case, the coupon rate on a per annum basis must be divided by 2 and the number of years to maturity must be multiplied by 2 before applying the equation, and the resulting yield must be multiplied by 2 to put it on a per annum basis. The formula as given above is not easy to solve for r when P, c, F, and n are given; however, there are prepared bond tables that can be used to look up the answer. See, for example, *Executive Bond Values Tables* (Boston: Financial Publishing Co., 1947).

[5] See J. R. Hicks, *Value and Capital* (2d ed.; Oxford: Clarendon Press, 1946), pp. 144–52; F. A. Lutz, "The Structure of Interest Rates," *Quarterly Journal of Economics*, Vol. LV, November 1940, pp. 36–63, reprinted in William Fellner and B. F. Haley (eds.), *Readings in the Theory of Income Distribution* (Philadelphia: Blakiston Co., 1946), pp. 499–529; J. W. Conard, *An Introduction to the Theory of Interest* (Berkeley: University of California Press, 1959), Part 3; B. G. Malkiel, *The Term Structure of Interest Rates: Expectations and Behavior Patterns* (Princeton, N.J.: Princeton University Press, 1966).

[6] Throughout this discussion, we use the capital letter R, with appropriate subscript, to indicate *actual yields* to maturity of securities as recorded in the market, and the small letter r, again with appropriate subscript, to indicate one-period interest rates which investors *expect* will prevail in future periods. Thus, for example, R_1 is the actual one-period rate currently prevailing, R_6 is the currently prevailing yield to maturity of a

(including accumulated interest) in each succeeding period are invested in one-period securities, the amount an investor will expect to have accumulated by the end of n periods is given by

$$(1 + R_1)(1 + r_2)(1 + r_3) \ldots . (1 + r_n) \qquad (1)$$

Let R_n be the prevailing yield to maturity on an n-period security. If a new security with a maturity of n periods is issued with a coupon rate R_n, this security will sell at par.[7] Now suppose an investor invests $1 in such a security, planning to hold it to maturity and to invest the interest coupons as he receives them at the successive short-term (one-period) rates which he expects to prevail over the life of the long-term security. The amount the investor would expect to accumulate by the end of n periods as a result of this procedure is given by the expression

$$R_n(1 + r_2)(1 + r_3) \ldots (1 + r_n) + R_n(1 + r_3)(1 + r_4) \ldots (1 + r_n) + \cdots + R_n(1 + r_n) + R_n + 1 \qquad (2)$$

The first term of this expression $[R_n(1 + r_2)(1 + r_3) \ldots (1 + r_n)]$ is the amount to which the first coupon payment (received at the end of period 1) is expected to accumulate by the end of n periods; the second term $[R_n(1 + r_3)(1 + r_4) \ldots (1 + r_n)]$ is the amount to which the second coupon is expected to accumulate; and so on, with the next to the last term (R_n) representing the last coupon to be received at the end of the nth period, and the last term representing the repayment of the $1 principal at the maturity of the security.

Since we are assuming that investors are intent on maximizing profits, we may suppose they will allocate their funds between one-period securities and n-period securities in such a way as to equate the expected returns on the two. That is, they will make the value of expression (2) above equal to that of expression (1). Thus, we have

$$R_n(1 + r_2)(1 + r_3) \ldots (1 + r_n) + R_n(1 + r_3)(1 + r_4) \ldots (1 + r_n) + \cdots + R_n(1 + r_n) + R_n + 1 = (1 + R_1)(1 + r_2) \ldots (1 + r_n) \qquad (3)$$

Since we are treating $R_1, r_2, \ldots, r_n$ as given, we can solve Equation 3 for R_n, as follows

$$R_n = \frac{(1 + R_1)(1 + r_2) \ldots (1 + r_n) - 1}{(1 + r_2)(1 + r_3) \ldots (1 + r_n) + (1 + r_3)(1 + r_4) \ldots (1 + r_n) + \cdots + (1 + r_n) + 1} \qquad (4)$$

---

security having five periods to run to maturity, r_2 is the one-period interest rate which investors expect will prevail in period 2, r_6 is the one-period interest rate which investors expect will prevail in period 6, and so on.

[7] It is immediately apparent from the equation given in footnote 4 that if the coupon rate on a security is equal to the prevailing yield to maturity on securities of that maturity (that is, if $c = r$), the security will sell at par (that is, $V = P$).

Equation 4 gives the equilibrium (profit-maximizing) value of R_n as a function of $R_1, r_2, r_3, \ldots, r_n$.

These ideas are illustrated in Table 14–1 which shows the structure of interest rates on bonds having maturities of from one to six years on a hypothetical date. In constructing the table, it was assumed that the current short-term (one-year) interest rate, R_1, was 3 percent and that the expected future one-year rates, r_2, r_3, r_4, r_5, and r_6, were 5 percent, 2 percent, 6 percent, 6 percent, and 8 percent, respectively. (This information is presented in the first three columns of the table.) The current and expected short-term rates are used to calculate the yields to maturity on bonds, R_2, R_3, R_4, R_5, and R_6, by means of Equation 4 above.[8] This set of yields, which constitutes the maturity structure of interest rates, is shown in column (6) of Table 14–1. The corresponding yield curve is shown in Figure 14–2.

TABLE 14–1

Hypothetical Structure of Interest Rates on Bonds Maturing in One to Six Years

Assumed Current and Expected One-Year Interest Rates			*Yields to Maturity*			
Year (1)	*Designation* (2)	*Value (Percent)* (3)	*Maturity (Years)* (4)	*Designation* (5)	*Yields Calculated from Equation (4) (Percent)* (6)	*Yields Calculated by Simple Average of One-Year Rates (Percent)* (7)
1	R_1	3.0	1	R_1	3.00	3.00
2	r_2	5.0	2	R_2	3.98	4.00
3	r_3	2.0	3	R_3	3.30	3.33
4	r_4	6.0	4	R_4	3.93	4.00
5	r_5	6.0	5	R_5	4.31	4.40
6	r_6	8.0	6	R_6	4.84	5.00

It should be noted that a reasonably good approximation to the results obtained by Equation 4 can be achieved by using the simple average of the current one-period rate and the expected future one-period rates over the relevant period to calculate the bond yields. The equation is

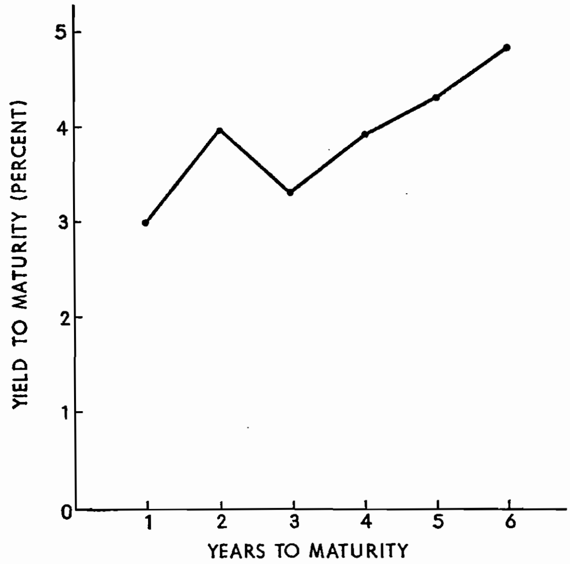
$$R_n = \frac{R_1 + r_2 + \cdots + r_n}{n} \tag{5}$$

[8] A typical calculation, that for R_3, is as follows

$$\begin{aligned} R_3 &= \frac{(1 + R_1)(1 + r_2)(1 + r_3) - 1}{(1 + r_2)(1 + r_3) + (1 + r_3) + 1} \\ &= \frac{(1.03)(1.05)(1.02) - 1}{(1.05)(1.02) + (1.02) + 1} \\ &= .03304 \end{aligned}$$

FIGURE 14–2

Hypothetical Yield Curve Constructed from Information in Table 14–1



The yields obtained by use of this approximation are given in column (7) of Table 14–1.[9]

[9] There is another equation that is often used to express the expectations hypothesis. It uses the following expression for the sum built up by investing for n periods in an n-period security

$$(1 + R_n)^n \tag{i}$$

This is then equated to expression (1) given in the text to obtain the following equation

$$(1 + R_n)^n = (1 + R_1)(1 + r_2)(1 + r_3) \ldots (1 + r_n)$$

Solving this equation explicitly for R_n, we obtain

$$R_n = [(1 + R_1)(1 + r_2)(1 + r_4) \ldots (1 + r_n)]^{\frac{1}{n}} - 1 \tag{ii}$$

The difference between this and Equation 4 given in the text arises from the fact that expression (i) above involves the assumption that the investor is able to invest the interest proceeds from the long-term investment each period at the long-term rate, R_n. This can readily be made apparent as follows: The sum built up at the end of n periods by investing \$1 at R_n and then reinvesting the interest at R_n can be expressed directly as follows:

$$R_n(1 + R_n)^{n-1} + R_n(1 + R_n)^{n-2} + \cdots + R_n(1 + R_n) + R_n + 1.$$

It can easily be shown by use of the formula for the sum of a geometric progression that the sum of all of the terms in the above expression prior to the last two is $(1 + R_n)^n - (1 + R_n)$. Thus, the expression simplifies to

$$(1 + R_n)^n - (1 + R_n) + R_n + 1 = (1 + R_n)^n$$

Since R_n is the long-term rate for n periods, it is unrealistic to assume that the investor would be able to reinvest the successive interest payments at this rate. Thus, Equation 4

If the relation given by Equation 4 holds, the expected return to the investor for a specified period will be the same regardless of the maturity of the security he chooses to invest in. To illustrate, suppose he invests in a newly issued one-year \$1,000 bond bearing a coupon rate of 3 percent. Since this bond will be selling at par, he will have to pay \$1,000 for it, and he will receive \$1,030 at the end of the year, thereby earning a return of exactly 3 percent.[10] Suppose instead he chooses to invest in a \$1,000 bond bearing a coupon rate of 4 percent and having three years to run to maturity. Since, according to column (6) of Table 14–1, the yield to maturity of three-year bonds is 3.30 percent, the price of a \$1,000 three-year bond bearing a 4 percent coupon can be calculated as follows:[11]

$$V = (.04)(\$1{,}000)\frac{1 - (1.0330)^{-3}}{.0330} + (\$1{,}000)(1.0330)^{-3}$$
$$V = \$1{,}018.85$$

One year later, this three-year security will have become a two-year security. Using the interest rate expectations shown in the first three columns of Table 14–1 in conjunction with Equation 4, the yield to maturity that is expected to prevail on two-year securities one year ahead can be calculated as follows:

$$(R_2)_{\text{Period 2}} = \frac{(1 + r_2)(1 + r_3) - 1}{(1 + r_3) + 1} = \frac{(1.05)(1.02) - 1}{1.02 + 1}$$
$$= .03515$$

That is, the yield on two-year securities one year later is expected to be 3.515 percent. The expected price of the \$1,000 4 percent bond corresponding to this rate can be calculated as follows:

$$V = (.04)(\$1{,}000)\frac{1 - (.03515)^{-2}}{.03515} + (\$1{,}000)(1.03515)^{-2}$$
$$V = \$1{,}009.21.$$

Thus, the expected return from buying the three-year bond and selling it one year later can be calculated as follows:

Purchase price	\$1,018.85	
Interest for 1 year	40.00	
Capital gain or loss (−) on sale of security (\$1,009.21 − \$1,018.85)	−9.64	
Net earnings for the year	\$ 30.36	= 3% of \$1,018.85

in the text, which assumes that the interest payments are reinvested at the sequence of *short-term* interest rates, is preferable. However, the results obtained with the two equations are generally quite similar; and Equation ii above is much less complex than Equation 4 in the text. Consequently, Equation ii has frequently been used—for example, in empirical studies of the term structure of interest rates.

[10] In these illustrations, it is assumed that bonds pay interest once a year.

[11] This is an application of the formula given in footnote 4 above.

Thus, the one-year "holding period yield" (including interest together with capital gain or loss) is 3 percent—the same as the current one-year interest rate.

This example illustrates a basic point: If the relation among interest rates expressed by Equation 4 holds true, an investor can expect the same return (including capital gain or loss) from investing funds for one year no matter what the maturity of the securities he invests in. And the same principle holds in general for funds invested for two years, three years, or any other period.

Some Qualifications

There are several reasons why the relationships explained above could hardly be expected to hold exactly.

1. All market participants do not have the same interest rate expectations, and these expectations are not held with complete confidence. Thus, the expectations that influence the interest rate structure are, in some sense, the "average" expectations of the various participants.

2. There are transactions costs involved in shifting funds from one maturity sector to another. These include not only explicit monetary costs of buying and selling securities but also the costs of keeping continuously informed about all prospective developments. Moreover, for some investors, the small prospective gains to be had from shifting funds from one maturity to another will not be judged to be worth the trouble involved. As a result of these factors, the market may show a certain amount of inertia, and the fine adjustments needed to achieve full equilibrium will not consistently be carried out.

3. Since expectations are subject to a considerable degree of uncertainty, some large institutional investors attempt to keep their funds invested in securities whose maturities are approximately the same as the maturities of their liabilities. Thus, life insurance companies typically show a market preference for long-term securities, commercial banks for short maturities, and so on. Such behavior may result in "market segmentation," which impedes the flow of funds among various maturities.

Some students of the maturity structure of interest rates have contended that these qualifications are so important as to undermine the validity of the expectations hypothesis. These observers commonly stress particularly the prevalence of market segmentation.[12] The more common view, however, is that despite these qualifications, expectations are the predominant force shaping the structure of rates. It should be noted that for expectations to be dominant, it is not necessary for all investors to be "arbitragers" who are

---

[12] See J. M. Culbertson, "The Term Structure of Interest Rates," *Quarterly Journal of Economics,* Vol. LXXI, November 1957, pp. 485–517.

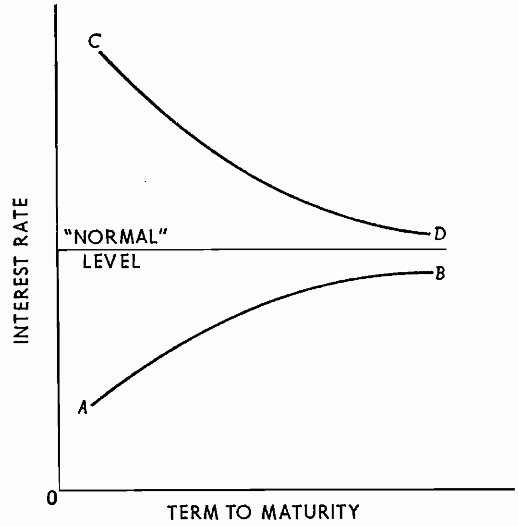
prepared to shift funds freely from one maturity sector to another in order to maximize profits. The behavior of some investors can be dominated by other factors—such as the desire to match the maturities of their assets and liabilities—provided there is a class of arbitragers who command a sufficient amount of funds to dominate the market.

The Nature of Expectations

If we assume that the expectations hypothesis holds true at least in a somewhat attenuated form, then the way in which the term structure of interest rates behaves will depend on the way in which expectations are formed and changed. The predominant hypothesis is that investors generally have *regressive interest rate expectations.* That is to say, at any particular time they have an opinion regarding the level of interest rates they regard as normal, and as short-term rates rise above or fall below this level, they expect them to regress back toward this normal level. Thus, as rates rise above normal, investors expect them to fall; and as rates fall below normal, investors expect them to rise. Using the simple average formulation (Equation 5) as an approximation of the relationship implied by the expectations hypothesis, it is apparent that (*a*) when short-term rates are expected to fall, current short-term rates will be above long-term rates and the yield curve will be negatively sloped, and (*b*) when rates are expected to rise, current short-term rates will be below long-term rates and the yield curve will be positively sloped. By the same reasoning, when the short-term rate is at approximately the level judged to be normal and is expected neither to rise nor to fall, rates for all maturities will be roughly equal and the yield curve will approximate a horizontal line. This would lead to patterns of rates such as those depicted in Figure 14–3. That is, when rates are high in relation to normal, they would be expected to fall and the yield curve would resemble *CD;* when rates are low and expected to rise, the yield curve would resemble *AB;* and when rates are not expected to change, the yield curve would be a horizontal line at the normal rate level.

It is apparent that the hypothetical patterns depicted in Figure 14–3 based on the expectations hypothesis with regressive expectations are broadly consistent with the actual rate patterns depicted in Figure 14–1. In March 1958, the economy was in recession, and interest rates were low by comparison with rates in the recent past as a result of the combined effects of the recession-weakened demand for credit and an easy monetary policy on the part of the Federal Reserve. Thus, the yield curve for March 1958 as shown in Figure 14–1 was upward-sloping, resembling the curve *AB* in Figure 14–3. On the other hand, in August 1959 and again in August 1969, the economy was experiencing inflationary pressures and interest rates were high in relation to the recent past as a result of strong credit demands and restrictive Federal Reserve policy. Hence the yield curves for August 1959

FIGURE 14–3
Hypothetical Patterns of Interest Rates



and August 1969 in Figure 14–1 were downward-sloping, resembling in a general way the curve *CD* in Figure 14–3.

There are two qualifications to the analysis presented above that need to be considered.

1. When interest rates are high and the yield curve is basically downward sloping, the curve often exhibits what has sometimes been called a "shoulder"—that is, the curve has a rising portion in the short maturity range, reaches a peak, and declines thereafter. This phenomenon is evident in the yield curves for August 1959 and August 1969 as depicted in Figure 14–1. The shoulder appears at about a three-year maturity for the August 1959 curve and at about a one-year maturity for the August 1969 curve. One possible explanation of the shoulder phenomenon is the existence of an element of what has been called *extrapolative expectations.* This means that when interest rates rise investors expect them to go on rising for a while before regressing toward their normal level.[13] While it seems probable that extrapolative expectations are sometimes present, it is doubtful whether their presence can fully explain the appearance of a shoulder of the kind that existed in August 1959. One problem is that at that time the extrapolative expectations hypothesis would imply that investors expected interest rates to continue to rise for about three years before beginning to decline, and

[13] See J. S. Duesenberry, *Business Cycles and Economic Growth* (New York: McGraw-Hill Book Co., 1958), pp. 318–19.

this is an implausibly long period. A second difficulty is that if extrapolative expectations are present in periods of high rates, they should sometimes manifest themselves in periods of low rates, leading to a dip in the yield curve in the short maturity range. There is no evidence of this. An alternative explanation of the shoulder is that there is a class of investors who have an inordinate preference for short-term securities. These investors are so anxious to avoid risk that they cannot be tempted to venture very far out along the yield curve during a period of high interest rates, even though there is a prospect of earning capital gains on longer term securities during the ensuing period of declining interest rates. Thus, these investors maintain a substantial demand for short-term securities even during periods of generally high interest rates, thereby supporting the prices of these securities and holding down their yields. Investors who behave in this way may include banks and nonfinancial corporations who hold U.S. government securities as a repository for their liquid funds.

2. The "normal" level toward which interest rates are expected to regress should not be viewed as a constant. Rather, it is probably in the nature of a moving average of past rates. Thus, if the short-term rate rises sharply, the yield curve is likely to take on a negative slope. However, if the rate remains at the higher level for a prolonged period, it will cause investors' ideas of the normal level of rates to be revised upward, and, as this happens, the rate curve will flatten. A decline in short-term rates that lasted for a prolonged period would have the opposite effect.

Concluding Comments

There is a considerable amount of empirical evidence in support of the expectations hypothesis as the basic explanation of the maturity structure of interest rates.[14] The most successful empirical models have generally made the interest rate on long-term Treasury bonds dependent on the relation between the three-month Treasury-bill rate and a moving average of past rates, in some cases attempting to take account of both regressive and extropolative expectations.[15] While these models have generally been fairly successful in explaining the maturity structure, there are still significant unsettled issues.

It seems doubtful whether the determinants of investor expectations can in all circumstances be satisfactorily captured by variables representing the

[14] See David Meiselman, *The Term Structure of Interest Rates* (Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1962); Frank de Leeuw, "A Model of Financial Behavior," in J. S. Duesenberry, Gary Fromm, L. R. Klein, and Edwin Kuh (eds.), *The Brookings Quarterly Econometric Model of the United States* (Chicago: Rand-McNally, 1965), pp. 464–530; Franco Modigliani and Richard Sutch, "Innovations in Interest Rate Policy," *American Economic Review,* Vol. LVI, May 1966, pp. 178–97.

[15] See, for example, Modigliani and Sutch, *op. cit.*

difference between current short-term rates and a moving average of past rates. There may be times when rates are very high by recent standards and would ordinarily be expected to regress toward normal but when there nevertheless exist circumstances which lead investors to expect a further rise. For example, investors may foresee large federal deficits which will require heavy Treasury borrowing, thereby exerting still further upward pressures on interest rates. Or inflationary pressures may be present, leading investors to expect a tightening of monetary policy by the Federal Reserve. A fully satisfactory explanation of interest-rate expectations is likely to have to incorporate "forward-looking" factors of this kind. This is, of course, very difficult to do because of the essentially subjective nature of the relevant expectational variables.

A final issue has to do with the effects of changes in the composition of the publicly held federal debt on the term-structure of interest rates. Such changes could be effected either by the Treasury in its management of the public debt, or by the Federal Reserve conducting open-market operations in different maturity sectors of the market. To pose the issue clearly: Would the maturity structure of interest rates on federal debt be significantly affected if the Federal Reserve were to sell $1 billion of 90-day Treasury bills out of its portfolio while simultaneously buying $1 billion of 25-year Treasury bonds? One's first commonsense reaction would be to say that this would raise the interest rate on Treasury bills while lowering the rate on long-term Treasury bonds. Indeed, such a response would be almost certain to occur immediately. But if the expectations hypothesis is correct, and if the Federal Reserve action did not change investor expectations, investors would act in such a way as to restore the preexisting rate structure, and the action would have no lasting effects. Several investigators have included variables designed to capture the maturity composition of the federal debt, or changes in that composition, in empirical models designed to explain the rate structure by means of expectations, in order to see if the addition of the debt-composition variables improves the explanatory power of the models.[16] In most cases, these efforts have met with little success, and the tentative conclusion seems to be that if changes in debt composition affect the term structure of interest rates, the effect is very weak or transitory.[17] This conclusion can, however, be regarded as only a tentative one. One difficulty may be that the Treasury itself responds to changes in the structure of interest rates in trying to achieve such objectives as minimizing the interest cost of

[16] This is done in de Leeuw, *op. cit.,* and Modigliani and Sutch, *op. cit.*

[17] For an empirical study which suggests that Treasury debt management and Federal Reserve actions can affect the maturity structure of interest rates, although not very powerfully, see A. M. Okun, "Monetary Policy, Debt Management, and Interest Rates: A Quantitative Appraisal," in *Stabilization Policies,* Research Studies Prepared for the Commission on Money and Credit (Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1963), pp. 331–80.

financing the debt, and it is therefore difficult to disentangle the effects of the term structure on the composition of the debt from the effects of the composition of the debt on the term structure.[18] Another difficulty may be that, in practice, the changes in the maturity composition of the debt have been too small to have much effect; larger changes in debt composition might have an appreciable influence on the rate structure.

Other Features of the Rate Structure

The above discussion suggests the existence of a fairly well determined and systematic relationship among interest rates on U.S. government securities of different maturities. If regressive expectations dominate (as usually seems to be the case) when interest rates move up and down, short-term rates will sweep over a wider range than long-term rates, as indicated in Figures 14–1 and 14–3.

Interest rates on various kinds of private debt are related through a process of investor arbitrage to interest rates on U.S. government securities. For example, at a given time the interest rate on bonds issued by a private corporation and having 10 years to run to maturity will lie above the interest rate on 10-year U.S. government bonds by an amount that is judged by the consensus of investors to provide sufficient compensation to cover the risk of default on the corporate bonds. If interest rates on U.S. government securities should fall, causing an abnormally large margin to arise between yields on such bonds and yields on corporate bonds, investors would purchase corporate bonds until their prices had risen and their yields had declined enough to restore the differential that investors judged appropriate in view of the risk.

Tax-exempt bonds issued by state and local governments, like corporate bonds, differ according to the degree of risk attributed to them by investors; and their yields move with the yields on U.S. government securities of equivalent maturity in much the same way as corporate bonds. However, their yields are lower than those on corporate bonds of equal quality as a result of the tax exemption. Furthermore, the tax exemption introduces an additional feature that may affect yield differentials between taxable and tax-exempt bonds. To illustrate, suppose at a given time the yield on a given quality of 15-year corporate bonds is 5 percent. If the quantity of tax-exempt bonds whose degree of risk is judged to be the same as the corporate bonds is such that all of the tax-exempt securities are held by taxpayers whose marginal tax rates under the federal income tax are 60 percent or higher, the yield on the tax-exempt bonds should be 2 percent. The reason for this is that the marginal taxpayer buying tax-exempt bonds will be able

[18] See W. L. Smith, "Comment," *American Economic Review*, Vol. LVI, May 1966, pp. 198–200.

to earn 2 percent after taxes [5 percent × (1 − .6) = 2 percent] on corporate bonds, so that he would benefit from buying tax-exempt bonds instead until their yield had fallen to 2 percent. Now suppose state and local governments issue large quantities of additional bonds to finance schools, highways, and other public projects; and in order to attract sufficient investment funds to absorb these bonds, they find it necessary to appeal to a much broader group of investors than originally with marginal tax rates as low as 30 percent. If the yield on equivalent corporate bonds remains at 5 percent, the yield on tax-exempt bonds will have to rise to 3.5 percent [5 percent × (1 − .3) = 3.5 percent] in order to equalize the yield on these bonds with the after-tax yield on corporate bonds for taxpayers in the 30 percent marginal tax bracket. Thus, the relation between yields on taxable and tax-exempt bonds depends on the volume of tax-exempt bonds, which determines how far down the tax scale it is necessary to reach to find buyers for the bonds.

In general, interest rates move up and down together as economic conditions and the monetary and fiscal policies of the government change. For the most part, through a process of arbitrage, interest rates on private debt move in such a way as to maintain reasonably stable relationships with interest rates on U.S. government securities of equivalent maturity.

Some types of interest rates are rather "sticky," however, and fail to move freely as the general level of rates changes. If permitted to do so, rates paid by banks on time and savings deposits might move fairly flexibly in line with the general level of short-term interest rates. This is because many savers are rather sensitive to interest-rate differentials and move their savings back and forth between deposits in the banks and short-term securities, depending on which will give them the higher return. This sensitivity makes bank deposits competitive with securities and works to keep the interest rates on the two in line. However, the banking authorities, for a variety of reasons, maintain ceilings on the interest rates the banks are permitted to pay on time and savings deposits and often do not permit these rates to rise in line with interest rates on securities. Rates charged by banks on many types of loans do not fluctuate as much as rates on open-market securities. This is partly due to the existence of usury laws which set an upper limit on rates banks can charge. In addition, the markets for some types of bank loans—notably loans to business—are often rather oligopolistic, involving rivalry among a relatively small group of banks; and interest rates on these loans exhibit the stickiness or rigidity that is characteristic of prices in oligopolistic markets generally. While mortgage interest rates move with interest rates on open-market securities, their range of fluctuation is commonly smaller. There are a number of reasons for this, including the existence of state usury laws and ceilings imposed by the federal government on the interest rates that can be charged on FHA-insured and VA-guaranteed mortgages.

DETERMINATION OF THE MONEY SUPPLY

In the earlier chapters of this book, we have treated the supply (i.e., stock) of money as though it were directly under the control of the central bank. Indeed, we have represented the money supply as a policy instrument that is capable of being precisely adjusted by the central bank in the conduct of monetary policy.

It is clearly a gross oversimplification to treat the money supply as being directly controlled by the central bank. The true instruments of monetary policy are primarily the central bank's portfolio of government securities, the discount rate it charges commercial banks when they borrow from it, and the cash reserve requirements it imposes on commercial banks.[19] The central bank cannot easily and directly control the money supply through the use of these instruments, since, as we shall see, the money supply is an endogenous variable that is determined jointly by the actions of the central bank, the responses of the commercial banks, and the public's preferences for different kinds of assets.[20]

A Model of the Financial Sector of the Economy

As a starting point for an analysis of the forces determining the money supply, it is useful to begin with a consideration of the sources and uses of bank reserves—that is, the factors supplying such reserves to the economy and the uses to which these reserves are put. The reserves of member banks[21]

---

[19] In addition to these primary instruments, the Federal Reserve System makes use of adjustable ceilings on the interest rates that commercial banks are permitted to pay on time and savings deposits as a supplemental instrument of monetary policy. Central banks in other countries make use of a variety of instruments that are not employed in the United States, including loan quotas for the commercial banks and direct regulation of the interest rates and other terms applicable to bank loans. Devices of this kind have been used to a limited extent in the United States, particularly during World War II and the Korean War. However, since they are not normally used here, we shall not consider them in this book.

[20] For a further elaboration of the analysis presented in this section, see W. L. Smith, "Time Deposits, Free Reserves, and Monetary Policy," in Giulio Pontecorvo, R. P. Shay, and A. G. Hart (eds.), *Issues in Banking and Monetary Analysis* (New York: Holt, Rinehart, & Winston, 1967), pp. 79–113.

[21] For a detailed explanation of the sources and uses of member bank reserves, which are derived by a consolidation of the balance sheet of the Federal Reserve System with a balance sheet covering certain monetary accounts of the U.S. Treasury, the student is referred to H. D. Hutchinson, *Money, Banking, and the United States Economy* (New York: Appleton-Century-Crofts, 1967), Chapter 7 including Appendix. A table showing the sources and uses of member bank reserves, under the heading, "Member Bank Reserves, Federal Reserve Bank Credit, and Related Items," is included at the beginning of the statistical section of the *Federal Reserve Bulletin* each month. See, for example, *Federal Reserve Bulletin,* December 1969, pp. A4 and A5. The item labelled "Other sources and uses, net," in Table 14–2, includes the following items from the table in the *Federal Reserve Bulletin:* float, other Federal Reserve assets, gold stock, Treasury cur-

of the Federal Reserve System consist of the deposits of these banks with their district Federal Reserve banks, together with cash held in the vaults of member banks. On the slightly simplifying assumption that all commercial banks are members of the Federal Reserve System, the sources and uses of reserves are summarized in Table 14–2.[22]

TABLE 14–2

Sources and Uses of Member Bank Reserves

Sources	*Uses*
P = Federal Reserve portfolio of U.S. government securities	$R_q{}^d$ = Required reserves for demand deposits
R_b = Bank borrowing from Federal Reserve	$R_q{}^t$ = Required reserves for time deposits
A = Other sources and uses, net	R_e = Excess reserves
	N = Currency outside banks

Reserves are affected by Federal Reserve purchases and sales of U.S. government securities which change its portfolio of such securities (designated by P in Table 14–2), by changes in bank borrowing at the Federal Reserve (R_b), and by changes in a variety of other factors lumped together under the heading "Other sources and uses, net" (A). The reserves generated through these forces are absorbed by becoming required reserves for demand deposits ($R_q{}^d$), by becoming required reserves for time deposits ($R_q{}^t$), by becoming excess reserves (R_e), or by being paid out to the public in the form of currency (N). The source and use tabulation presented in Table 14–2 is an accounting statement of such a nature that the sum of the sources must always be equal to the sum of the uses. That is, it must be true under all circumstances that

$$P + R_b + A = R_q{}^d + R_q{}^t + R_e + N.$$

rency outstanding, Treasury cash holdings, Treasury deposits with Federal Reserve Banks, foreign deposits with Federal Reserve Banks, other deposits (excluding those of member banks) with Federal Reserve Banks, and other Federal Reserve liabilities and capital.

[22] Since all banks are not members of the Federal Reserve System, it is necessary in practice to treat Currency in Circulation (which includes all currency and coin outside the Treasury and the Federal Reserve System) as a use of member bank reserves. The resulting table presents the sources and uses of that portion of member bank reserves held in the form of deposits by member banks with the Federal Reserve banks. To account for all reserves, it is then necessary to add in the amount of vault cash held by member banks. If all banks were members of the Federal Reserve System, it would be possible to present a sources and uses analysis covering all member bank reserves including vault cash, as is done in Table 14–2, by treating Currency outside Banks (which includes all currency and coin held outside the Treasury, the Federal Reserve System, and the commercial banks) as a use instead of only Currency in Circulation. Since Currency outside Banks is the component for currency (and coin) that is included along with demand deposits held by the public in computing the money supply, the assumption that all banks are members of the Federal Reserve System results in a considerable simplification.

This identity can be rewritten as

$$P + A = R_q^d + R_q^t + R_e - R_b + N \tag{6}$$

In this form, it says that $P + A =$ Unborrowed reserves $(R_q^d + R_q^t + R_e - R_b)$ plus currency (N). Thus, by controlling $P + A$, the Federal Reserve can determine unborrowed reserves plus currency. It is able to control $P + A$ quite accurately by means of open-market purchases and sales of U.S. government securities. Employing a distinction that has frequently been made, we can say that by engaging in "defensive" open-market operations to change P (which it can control exactly) to offset undesired changes in A (which it cannot control), it can prevent undesired changes in $P + A$. On the other hand, it can use "dynamic" open-market operations to change P in such a way as to produce the changes it wants in $P + A$ for the purpose of conducting monetary policy to promote economic stability.[23] We shall substitute

$$W = P + A \tag{7}$$

into Equation 6, W (which is equal to unborrowed reserves plus currency) being the magnitude which we shall assume the Federal Reserve controls by open-market operations.

In analyzing U.S. monetary policy, considerable use is made of the concept of "free reserves," which is the difference between excess reserves and bank borrowing from the Federal Reserve. If we designate free reserves as F, we have

$$F = R_e - R_b \tag{8}$$

When monetary policy is relatively easy, F is commonly a positive number —that is, excess reserves are greater than bank borrowing from the Federal Reserve. On the other hand, in periods when monetary policy is restrictive, bank borrowing is frequently greater than excess reserves and F is a negative number.

Substituting Equations 7 and 8 into Equation 6 and rearranging terms, we have

$$W - F = R_q^d + R_q^t + N \tag{9}$$

If D is the amount of demand deposits, T is the amount of time deposits, and d and t are the reserve requirements applicable to demand deposits and time deposits, respectively, we may write

$$R_q^d = dD \tag{10}$$
$$R_q^t = tT \tag{11}$$

[23] The distinction between "defensive" and "dynamic" open-market operations is developed in R. V. Roosa, *Federal Reserve Operations in the Money and Government Securities Markets* (New York: Federal Reserve Bank of New York, 1956).

Substituting Equations 10 and 11 into Equation 9, we obtain

$$W - F = dD + tT + N \tag{12}$$

Despite the manipulations we have performed, Equation 12 is still basically an accounting identity derived from Table 14–2.

Now we define the money supply (M) as the sum of demand deposits and currency; that is

$$M = D + N.$$

Let us suppose the public follows the practice of holding a fixed proportion, v, of the money supply in the form of currency (and the remainder, of course, in demand deposits).[24] That is

$$N = vM \tag{13}$$

$$D = (1 - v)M \tag{14}$$

where v is a positive fraction less than unity. Substituting Equations 13 and 14 into Equation 12, we have

$$W - F = [d(1 - v) + v]M + tT.$$

If we let $m = d(1 - v) + v$, we can rewrite this as

$$W - F = mM + tT \tag{15}$$

The coefficient m is the effective reserve requirement for money. It is equal to the reserve requirement for demand deposits (d) multiplied by the proportion of the money supply that is in the form of demand deposits, plus the reserve requirement for currency (unity) multiplied by the proportion of the money supply that is in the form of currency.

There is considerable evidence that the demand for free reserves by the banks depends on market interest rates and on the Federal Reserve discount rate. In linear form the equation for the demand for free reserves might be as follows

$$F = ar_d + br + F_0 \tag{16}$$

where r_d is the Federal Reserve discount rate, r is the short-term market interest rate (on 90-day Treasury bills, for example), and F_0 is a constant term. We shall assume, in accordance with the empirical evidence, that

[24] Instead of assuming that currency and demand deposits are held in constant proportions, it would be more accurate to employ separate demand equations for each, since there is considerable evidence that they respond differently to changes in income and interest rates. On this, see R. L. Teigen, "The Demand for and Supply of Money," in W. L. Smith and R. L. Teigen (eds.), *Readings in Money, National Income, and Stabilization Policy* (rev. ed.; Homewood, Ill.: Richard D. Irwin, Inc., 1970). It would be quite possible to expand the model we are employing to handle currency and demand deposits in this more realistic way. However, this would increase the complexity of the presentation considerably without altering the conclusions in any fundamental way.

$a > 0$ and $b < 0$. When the discount rate is raised with market interest rates constant, free reserves increase as the banks reduce their borrowing from the Federal Reserve and increase their holdings of excess reserves in order to reduce the chance that they will be forced to borrow. Conversely, when market interest rates rise with the discount rate constant, they increase their borrowing and reduce their excess reserves.[25]

There is abundant evidence, as explained in Chapter 11, that the demand for money depends on income and interest rates; in addition, we shall assume that it depends on wealth. Our demand for money equation is

$$M = cY + eK + hr + jp \tag{17}$$

where Y is GNP, K is the stock of wealth of the community, r is the interest rate on short-term securities, and p is the interest rate paid by the banks on time deposits. We assume that $c > 0$, $e > 0$, $h < 0$, and $j < 0$. That is, the public attempts to economize the use of money when either the interest rate on securities (r) or the interest rate on time deposits (p) rises.

We shall assume the existence of a demand for time deposits, which, like that for money, depends on income, wealth, and interest rates. The demand equation for time deposits is

$$T = fY + gK + kr + wp \tag{18}$$

In this case, $f > 0$, $g > 0$, $k < 0$, and $w > 0$. That is, the higher the interest rate on securities, the less time deposits the public will hold, while the higher the interest rate on time deposits, the more of them will be held.[26]

Since a rise in the interest rate on securities will cause the public to shift funds out of time deposits into the now higher yielding securities (this is implied by the fact that $k < 0$), we may suppose the banks will attempt to avoid the resulting loss of time deposits by raising interest rates on such deposits in response to increases in the interest rate on securities, although not necessarily by the full amount of the increase in the interest rate on securities. Thus, we have an equation

$$p = nr + p_0 \tag{19}$$

where p_0 is a constant term and $1 \geq n \geq 0$. This is a form of supply equation for time deposits—that is, it implies that the banks set the interest rate on time deposits in relation to the interest rate on securities and then accept all the funds that are offered for deposit at that rate.

Since changes in interest rates affect spending and income with a con-

[25] For a discussion of the relation of free reserves to the money supply, see Teigen, *op. cit.* On free reserves and their behavior generally, see A. J. Meigs, *Free Reserves and the Money Supply* (Chicago: University of Chicago Press, 1962); W. L. Smith, "The Instruments of General Monetary Control," *National Banking Review*, Vol. I, September 1963, pp. 47–76, reprinted in Smith and Teigen, *Readings*, rev. ed., *op. cit.*

[26] On the demand for time deposits and its relation to the demand for money, see Teigen, *op. cit.*

siderable lag, and since changes in the community's wealth occur only gradually as saving leads to wealth accumulation, we shall, for purposes of this analysis, treat income and wealth as constants; thus, we will set Y equal to a constant value $\bar{Y}$, and K equal to a constant value $\bar{K}$ in Equations 17 and 18. This permits us to isolate the financial sector of the economy, leaving out the links from the financial to the real sector and the feedbacks from the real to the financial sector. This is a legitimate form of partial equilibrium analysis, although a more complete analysis would require us to take the links and feedbacks into account.

Analysis of the Model

The model we are using is composed of the following five equations, which are repeated for convenience:

$$W - F = mM + tT \tag{15}$$
$$F = ar_d + br + F_0 \tag{16}$$
$$M = c\bar{Y} + e\bar{K} + hr + jp \tag{17}$$
$$T = f\bar{Y} + g\bar{K} + kr + wp \tag{18}$$
$$p = nr + p_0 \tag{19}$$

This model contains five endogenous variables: the amount of free reserves (F), the stock of money (M) which is composed of currency and demand deposits in constant proportions, the stock of time deposits (T), the interest rate on securities (r), and the interest rate on time deposits (p). There are four policy instruments that the central bank may employ in conducting monetary policy: W, which represents open-market operations; r_d, the discount rate; d, the reserve requirement for demand deposits; and t, the reserve requirement for time deposits. [Since with a constant ratio, v, of currency to the money stock, we have $m = d(1 - v) + v$, a change in d will affect m in Equation 15]. We shall consider only the effects of open-market operations, but the other instruments could be analyzed in a similar way. The model analyzes the interacting markets for four financial assets, one held by the banks (free reserves), one held by both the banks and the public (securities), and two held by the public (money and time deposits).[27] It is useful to think of the securities in the model as short-term securities, such as 90-day Treasury bills which are close substitutes for money and time deposits in the portfolios of investors. Thus, the interest rate, r, should be viewed as the short-term interest rate.

[27] In a model such as this one containing four markets, if three of the markets are equilibrated, the fourth must also be in equilibrium. Thus, it is not necessary to include specific equations for the fourth market. In this model, there are, in effect, supply and demand equations for free reserves, money, and time deposits; it is not necessary to include equations for the market for securities. See footnote 5 in Chapter 12 above.

To analyze the working of the model, we begin by substituting Equations 16, 17, 18, and 19 into Equation 15. This process yields

$$W - ar_d - br - F_0 = mc\bar{Y} + me\bar{K} + mhr + mjnr + mjp_0 + tf\bar{Y} + tg\bar{K} + tkr + twnr + twp_0 \qquad (20)$$

Solving Equation 20 for r, we obtain

$$r = \frac{1}{mh + mjn + tk + twn + b}[W - ar_d - F_0 - (mc + tf)\bar{Y} - (me + tg)\bar{K} - (mj + tw)p_0] \qquad (21)$$

In the context of this model, which contains only three financial assets—money, time deposits, and securities—that are held by the public, any funds that are shifted into time deposits as a result of a rise in the interest rate on time deposits must come either from money or from securities. If it is assumed that a given rise in the interest rate on time deposits will induce the shifting of the same amount of funds from securities to time deposits as would be induced by an equal decline in the interest rate on securities, we can write

$$w = -j - k \qquad (22)$$

In the case of a rise in the interest rate on time deposits, w measures the total volume of funds attracted into time deposits, $-j$, measures the portion of these funds derived from a shift out of money balances, and $-k$ measures the portion derived from a shift out of securities.

Substituting Equation 22 into Equation 21, we obtain

$$r = \frac{1}{mh + (m - t)nj + t(1 - n)k + b}[W - ar_d - F_0 - (mc + tf)\bar{Y} - (me + tg)\bar{K} - (mj + tw)p_0] \qquad (23)$$

Next, substituting Equation 19 into Equation 17, we have

$$M = (h + nj)r + c\bar{Y} + e\bar{K} + jp_0 \qquad (24)$$

Substituting Equation 23 for r in Equation 24, we obtain

$$M = \frac{h + nj}{mh + (m - t)nj + t(1 - n)k + b}[W - ar_d - F_0 - (mc + tf)\bar{Y} - (me + tg)\bar{K} - (mj + tw)p_0] + c\bar{Y} + e\bar{K} + jp_0 \qquad (25)$$

To find the effects of a change in the Federal Reserve System's portfolio of U.S. government securities on the interest rate on securities and on the stock of money, take the first differences of Equations 23 and 25, holding all of the variables in brackets constant except W. This yields

$$\Delta r = \frac{1}{mh + (m - t)nj + t(1 - n)k + b}\Delta W \qquad (26)$$

$$\Delta M = \frac{h+nj}{mh + (m-t)nj + t(1-n)k + b} \Delta W \qquad (27)$$

A Numerical Illustration

Perhaps the easiest way to understand how this model works is to take a numerical example. Suppose the parameters and constants of the model have the following values: $d = .20$, $v = .25$, $m(= d(1-v) + v) = .40$, $t = .05$, $a = .55$, $b = -.55$, $F_0 = 0$, $c = .15$, $e = .05$, $h = -5$, $j = -5$, $f = .10$, $g = .10$, $k = -20$, $w(= -j - k) = 25$, $n = .6$, $p_0 = 0$, $\bar{Y} = 1{,}000$, $\bar{K} = 2{,}000$, $r_a = 6$, $W = 94.3$. Then Equations 15–19 become

$$W - F = .40M + .05T \qquad (15a)$$
$$F = .55r_a - .55r \qquad (16a)$$
$$M = .15\bar{Y} + .05\bar{K} - 5r - 5p \qquad (17a)$$
$$T = .10\bar{Y} + .10\bar{K} - 20r + 25p \qquad (18a)$$
$$p = .6r \qquad (19a)$$

Substituting the values of the appropriate parameters and constants into Equation 23 we obtain

$$r = 6$$

The values of the other endogenous variables are as shown in the "Original Equilibrium" column of Table 14–3.

TABLE 14–3

Effects of Open-Market Operations on Financial Variables
(dollar amounts in billions)

Variable	*How Calculated*	*Original Equilibrium*	*New Equilibrium**	*Change*
W	Exogenous	\$ 94.30	\$ 98.30	+\$4.00
F	$.55r_a - .55r$	0	0.55	+ 0.55
$W - F$	$.40M + .05T$	94.30	97.75	+ 3.45
M	$.15\bar{Y} + .05\bar{K} - 5r - 5p$	202.00	210.00	+ 8.00
T	$.10\bar{Y} + .10\bar{K} - 20r + 25p$	270.00	275.00	+ 5.00
r_a	Exogenous	6.0%	6.0%	...
r	From Equation 23	6.0%	5.0%	−1.0%
p	$.6r$	3.6%	3.0%	−0.6%

* After open-market purchase of \$4 billion of U.S. government securities by the Federal Reserve System.

Now suppose the Federal Reserve purchases \$4 billion of U.S. government securities in the open market. Substituting the values of the relevant parameters in Equation 26, we obtain

$$\Delta r = -.25\Delta W$$

For the indicated $4 billion of open-market purchases, this becomes

$$\Delta r = -.25(4) = -1.0 \text{ percent.}$$

That is, $4 billion of open-market purchases will lower the interest rate on securities from 6 percent to 5 percent.

To find the effect on the money stock we substitute the relevant parameter values into Equation 27, obtaining

$$\Delta M = 2.0\Delta W = 2.0(4) = \$8.0 \text{ billion}$$

That is, the $4 billion of open market purchases will increase the money supply from $202.0 billion to $210.0 billion. The value of all the endogenous variables after the open-market purchase of $4 billion are shown in the "New Equilibrium" column of Table 14–3.

It will be useful to attempt to sort out the effects of the open-market purchase a little more carefully. The effects of the operation on the balance sheets of the nonbank public, the commercial banks, and the Federal Reserve System are shown in Table 14–4.

The original purchase of $4 billion of securities by the Federal Reserve adds $4 billion to the cash reserves of the commercial banks. The entries on

TABLE 14–4

Effects of Open-Market Purchase of $4 Billion on Balance Sheets of Nonbank Public, Commercial Banks and Federal Reserve
(amounts in billions of dollars)

Nonbank Public

Assets		Liabilities	
Demand deposits	+ 6.00		
Time deposits	+ 5.00		
Currency	+ 2.00		
Securities	−13.00		

Commercial Banking System

Assets		Liabilities	
Securities*	− 4.00	Borrowing from Federal Reserve	−0.55
Cash reserves*	+ 4.00	Demand deposits	+6.00
Cash reserves	− 0.55	Time deposits	+5.00
Securities	+13.00		
Cash reserves	− 2.00		

Federal Reserve System

Assets		Liabilities	
Securities*	+ 4.00	Deposits of commercial banks (reserves)*	+4.00
Lending to commercial banks	− 0.55	Deposits of commercial banks (reserves)	−0.55

* Entries for original purchase of securities by the Federal Reserve.

the balance sheets of the commercial banks and the Federal Reserve System for this initial purchase are marked with asterisks in Table 14–4 to separate them from the remaining entries.[28] This purchase lowers interest rates to some degree because the Federal Reserve must pay a high enough price for securities to induce the banks to part with them. Moreover, the increased reserves will enable the banks to purchase additional securities in the process of expanding credit. When the whole operation is completed, the interest rate on securities (r) will have fallen by 1 percent, and the interest rate on time deposits (p) will have fallen by 0.6 percent. These changes in interest rates will have caused the following adjustments.

1. In accordance with Equation 16a, a decline of 1 percent in the interest rate on securities with the Federal Reserve discount rate unchanged causes the commercial banks to increase their free reserves by $550 million. For simplicity, it is assumed that this entire adjustment occurs through a reduction of borrowing by the commercial banks from the Federal Reserve, although a portion of it would in practice presumably take the form of an increase in excess reserves.

2. The change in interest rates causes the public to make several adjustments in its portfolio of financial assets. In order to disentangle these effects, it will be useful to repeat Equations 17a and 18a

$$M = .15\bar{Y} + .05\bar{K} - 5r - 5p \qquad (17a)$$

$$T = .10\bar{Y} + .10\bar{K} - 20r + 25p \qquad (18a)$$

With the aid of these equations, the portfolio adjustments may be sorted out as follows:

a) Since the interest rate on securities has fallen by 1 percent while the interest rate on money is unchanged (at a value of zero), the public shifts funds from securities to money. The magnitude of this effect is measured by the term "$-5r$" in Equation 17a—since $\Delta r = -1$ percent, the shift from securities to money is $(-5) \times (-1) =$ \$5 billion.

b) Since the interest rate on time deposits has fallen by 0.6 percent while the interest rate on money is unchanged, the public shifts funds from time deposits to money. The size of this shift is measured by the term "$-5p$" in Equation 17a—since $\Delta p = -0.6$, the shift into money from time deposits is $(-5) \times (-0.6) =$ \$3 billion. Alternatively, it is reflected in a portion equal to $+5p$ of the term "$+25p$" in Equation 18a—that is, $(+5) \times (-0.6) = -$\$3 billion is the amount of the shift *out of* time deposits into money.

c) Since the interest rate on securities has fallen by 1 percent while the interest rate on time deposits has fallen by 0.6 percent, the yield differential

[28] It is assumed in this illustration that the commercial banks are the sellers of the securities purchased by the Federal Reserve. If the sellers were the nonbank public, the final outcome would be the same, although the detailed entries would be slightly different.

between securities and time deposits has narrowed by 0.4 percent, thereby making time deposits more attractive and causing the public to shift funds from securities to time deposits. As indicated above, $+5p$ of the term "$+25p$" in Equation 18a was "used up" in explaining the shift of funds from time deposits into money. The remaining interest rate effects in Equation 18a are reflected in the terms "$-20r + 20p$," which can be used to calculate the size of the shift of funds into time deposits from securities: $(-20) \times (-1.0) + (20) \times (-0.6) = (20) \times (0.4) =$ \$8 billion. These effects are summarized in Table 14–4.

3. As a result of the adjustments described above, the public reduces its holdings of securities by \$13 billion while increasing its holdings of money by \$8 billion and increasing its holdings of time deposits by \$5 billion, as shown in Table 14–5. According to Table 14–4, the commercial banks ex-

TABLE 14–5

Changes in Portfolios of the Nonbank Public Caused by Open-Market Purchase of \$4 Billion (amounts in billions of dollars)

		Effects on:		
	How Calculated	*Money*	*Time Deposits*	*Securities*
Shift from securities to money	$-5r$ in equation (17a)	+5		−5
Shift from time deposits to money	$-5p$ in equation (17a) or $+5p$ in equation (18a)	+3	−3	
Shift from securities to time deposits	$-20r + 20p$ in equation (18a)		+8	−8
Totals		+8	+5	−13

pand their demand deposits by \$6 billion and pay out \$2 billion of currency to the public (thereby expanding the money supply by \$8 billion) while at the same time experiencing an increase of \$5 billion in their time deposits. This enables the banks to acquire the \$13 billion of securities that are sold by the public. The banks' reserves were initially increased by \$4 billion as a result of the Federal Reserve's open-market purchases of securities. The banks use up \$0.55 billion of these reserves by repaying borrowings at the Federal Reserve and pay out \$2 billion to meet the public's increased demand for currency. Thus the banks wind up with \$1.45 billion more reserves than originally (\$4 billion − \$0.55 billion − \$2 billion = \$1.45 billion); this is precisely the amount of reserves needed as required reserves for their new demand and time deposits (\$6 billion × .20 + \$5 billion × .05 = \$1.45 billion).

Conclusions

Three main conclusions may be drawn from this analysis.

1. ***The Money Supply Is an Endogenous Variable of the System.*** Its value is jointly determined by (*a*) the monetary policy actions of the Federal Reserve as reflected in open-market purchases and sales of securities, as well as adjustments in the discount rate and in reserve requirements; (*b*) the actions of the commercial banks in adjusting their holdings of excess reserves and their borrowing from the Federal Reserve in response to changes in market interest rates; and (*c*) the actions of the nonbank public in adjusting the composition of its portfolios of money, time deposits, and securities in response to changes in interest rates.

It may be noted that in the standard Keynesian analysis of money and interest rates employed in Chapters 11–13 of this book, the only phenomenon allowed for among those analyzed above is the tendency of the public to adjust its demand for money in response to changes in the interest rate on securities. Thus, the Keynesian model assumes, in effect, that all of the interest rate coefficients in Equations 15–19 except h are equal to zero. If this is the case, Equation 27 reduces to

$$\begin{aligned}\Delta M &= \frac{1}{m}\,\Delta W \\ &= \frac{1}{d(1-v)+v}\,\Delta W\end{aligned}$$

This is a standard text book formulation of the so-called credit expansion multiplier for the banking system, which is alleged to establish an essentially mechanical relation between Federal Reserve open-market operations and the money supply.

If such a mechanical relation existed, it would be a relatively simple matter for the Federal Reserve to control the money supply almost exactly, and the money supply could, for all practical purposes, be regarded as the instrument through which the Federal Reserve conducted monetary policy. Even under a more complex setup of the kind expressed in Equations 15–19, it would, of course, still be possible for the Federal Reserve to control the money supply precisely, provided it knew the exact structure of the system and the structure did not change. In reality, however, the system is a good deal more complex than that reflected in Equations 15–19 because there are many more financial assets and correspondingly more interest rates involved than are taken account of in those equations. Moreover, we do not know the exact structure of the system at any given time, and there is every reason to believe that the structure is subject to frequent changes. Consequently, it is not easy for the Federal Reserve to control the money supply exactly because it is affected by a variety of shifting forces outside the reach of the Federal Reserve.[29]

2. ***All of the Effects We Have Discussed Act to Reduce the Effect of Open-Market Operations on Interest Rates.*** In addition to the standard Keynesian assumption that investors regard money and securities as substitutes in their portfolios, we have allowed for portfolio substitution between money and time deposits and between time deposits and securities and for the practice by banks of adjusting their free reserve position in response to changes in interest rates. If there is no substitution between money and time deposits, the coefficient j in Equation 17 will be zero and the second term $[(m - t)nj]$ in the denominator of Equation 26 will be reduced to zero. If there is no substitution between time deposits and securities, the coefficient k in Equation 18 will be zero and the third term $[t(1 - n)k]$ in the denominator of Equation 26 will be reduced to zero. If the banks' demand for free reserves is unresponsive to the interest rate on securities, the coefficient b in Equation 16 will be zero and the fourth term (b) in the denominator of Equation 26 will be reduced to zero. Since each of the terms in the denominator of Equation 26 is negative, the power of open-market purchases (sales) to lower (raise) the interest rate on securities will be increased by the reduction of any of these terms to zero.

If all three of these effects were absent so that all that remained was the simple Keynesian substitution between money and securities, Equation 26 would reduce to

$$\Delta r = \frac{1}{mh} \Delta W$$

In our numerical example, with $m = .40$ and $h = -5$, this would become

$$\Delta r = -.5\Delta W$$

That is, $1 billion of open-market purchases would lower the interest rate on securities by one-half percentage point as compared with only one-quarter percentage point in our example with all four effects present.

We may conclude that the effects we have been discussing all work to reduce the potency of monetary policy, at least as measured by the magnitude of the effects on interest rates produced by a given amount of monetary policy action. The reason for this, as the reader can no doubt discern for himself by careful thought, is that each of these effects increases the combined demand for securities by the banks and the public above what it would be in the absence of that effect.

3. ***The Money Supply Possesses a Degree of Interest Elasticity.*** It was assumed in earlier chapters that while the *demand* for money was elastic

[29] On the problems of controlling the money supply, see A. R. Holmes, "Operational Constraints on the Stabilization of Money Supply Growth," and S. J. Maisel, "Controlling Monetary Aggregates," in *Controlling Monetary Aggregates* (Boston: Federal Reserve Bank of Boston, 1969), pp. 65–77 and 152–74, respectively.

to the interest rate, the *supply* of money was completely interest inelastic. When substitution between money and time deposits and between time deposits and securities and the response of free reserves to interest rates are taken into account, however, the money supply is almost certain to be interest elastic.

A money supply function can be derived from Equations 15–19 above. Upon substituting Equations 16 and 18 into Equation 15 and solving for M, we obtain

$$M = \frac{1}{m}\,[W - ar_d - br - F_0 - tf\bar{Y} - tg\bar{K} - tkr - twp]$$

Using Equation 19 to eliminate p, this can be reduced to

$$M = -\,\frac{t(k + nw) + b}{m}\,r + \frac{1}{m}\,(W - ar_d - F_0 - twp_0 - tf\bar{Y} - tg\bar{K})$$

Using the condition that $w = -j - k$ (Equation 22), this becomes

$$M = -\,\frac{t((1 - n)k - nj) + b_r}{m} + \frac{1}{m}\,(W - ar_d - F_0 + t(j + k)p_0 - tf\bar{Y} - tg\bar{K}) \qquad (28)$$

Substituting the values from the numerical example employed earlier into Equation 28, we obtain

$$M = 2r + 190 \qquad (29)$$

This is the money supply equation for the model depicted by Equations 15a–19a when $W = \$94.3$ billion, $r_d = 6$ percent, $\bar{Y} = \$1{,}000$ billion, and $\bar{K} = \$2{,}000$ billion. The corresponding money demand equation, obtained by substituting Equation 19a into Equation 17a is

$$M = -8r + 250 \qquad (30)$$

Solving Equation 29 and 30 simultaneously, we obtain the values $M = \$202$ billion and $r = 6$ percent which were given earlier (see Table 14–3).

According to Equation 29, the money supply responds posivitely to changes in the interest rate—other things remaining constant, a rise of one percentage point in the interest rate will bring forth $2 billion of additional money. There are two forces that tend to cause such a positive response of the money supply to the interest rate. First, the rise in the interest rate causes the banks to reduce their free reserves—that is, to increase their borrowings from the Federal Reserve and to reduce their excess reserves—and this provides the basis for an expansion of bank credit and the money supply. Second, if the interest rate on time deposits is not fully adjusted to the interest rate on securities (i.e., if $n < 1$), investors will shift funds from time deposits to securities. Those who sell the securities to the former holders of time deposits will retain the proceeds in the form of additional money balances, and this will cause the money supply to increase.[30] Thus, the shift of funds from time deposits to securities has indirectly caused a shift from time deposits to money. Since money bears higher reserve requirements than time deposits, some secondary contraction of security holdings and money by the banks will be necessary; but this will not be great enough to wipe out the initial increase in the money supply. There is a third factor which will tend to cause the money supply to contract as the interest rate rises. If the interest rate on time deposits rises at all in response to the rise in the interest rate on securities (i.e., if $n > 0$), some funds are likely to be shifted from money (whose interest rate is fixed at zero) into time deposits, and this shift will reduce the money supply. It is conceivable but unlikely that this third negative factor will outweigh the two positive ones, leading to a situation in which the money supply declines as the interest rate rises.

AN EXPANDED MODEL

It will be useful to expand the model developed in earlier chapters of this book to see how some of the ideas presented in this chapter can be fitted in. The model we shall employ consists of the following equations.

$$C = cY_d + C_0 \quad (31)$$
$$I = iY_d - vr + I_0 \quad (32)$$
$$Y_d = Y - T \quad (33)$$
$$T = xY + T^* \quad (34)$$
$$Y = C + I + G^* \quad (35)$$
$$M_d = kY - ms + L_0 \quad (36)$$
$$M_s = hW^* + qs - pr_d^* \quad (37)$$
$$M_d = M_s \quad (38)$$
$$r - \bar{r} = w(s - \bar{r}) \quad (39)$$

Here C is personal consumption expenditures, Y_d is disposable (after-tax) income, I is private investment expenditures, r is the long-term interest rate, Y is GNP, T is net taxes (taxes minus transfer payments), G^* is government purchases of goods and services, M_d is the quantity of money demanded, M_s is the quantity of money supplied, s is the short-term interest rate, W^* is unborrowed reserves plus currency, and r_d^* is the central bank's discount rate. The model contains nine equations and the following nine endogenous variables: $C, Y_d, I, r, Y, T, M_d, M_s$, and s. There are four policy instruments: unborrowed reserves plus currency (W^*), which is assumed to be controlled by the central bank through open market operations; the discount rate of

[30] The recipients must hold the proceeds in money form; if they put the funds into time deposits, there will not have been the posited shift of funds from time deposits to securities for the system as a whole.

the central bank (r_d*); government purchases of goods and services (G*); and the level of taxes (T*). However, we shall discuss the use of only two of these instruments, W* and G*.

In most respects this model is familiar, since Equations 31–35 are exactly the same as the corresponding equations in the model employed in Chapter 13. The innovations are to be found in the last four equations, and they consist basically of three.

1. The money supply is an endogenous variable, the determinants of which are depicted in Equation 37. This equation, which is an approximation to Equation 28 derived from the model presented in the previous section of this chapter, indicates that the money supply (*a*) increases when the central bank's portfolio of government securities (G*) is increased; (*b*) increases when the short term interest rate (s) increases; and (*c*) decreases when the central bank's discount rate (r_d*) is increased. The dependence of the money supply on the short-term market interest rate reflects the responses discussed earlier in this chapter—the tendency for the banks to increase their borrowing from the central bank when interest rates rise, and the tendency of the public to adjust the composition of their holdings of financial assets (money, time deposits, and securities) when interest rates change.[31]

2. A distinction is made between the short-term and long-term interest rates, and it is assumed that the short-term interest rate is determined in the money market—that is, by the interaction of the demand for money and the supply of money (Equations 36, 37, and 38).

3. Equation 39 depicts the behavior of the maturity structure of interest rates. The parameter $\bar{r}$ may be viewed as the "normal" level of interest rates—that is, the level toward which investors expect rates to regress when they depart from that level. If the coefficient w is positive and less than unity, the long-term interest rate will move in the same direction as the short-term rate but by a smaller amount. Moreover, when rates are at their "normal" level, the short-term rate will be equal to the long-term rate. Suppose, for example, that $\bar{r} = 5$ percent and $w = .25$ so that Equation 39 becomes

$$r - 5 = .25(s - 5)$$

[31] It was explained at the end of the previous section of this chapter that there are two forces that tend to cause the money supply to increase when the interest rate rises and one force that tends to cause it to fall. In the present model we are assuming that the two positive forces outweigh the negative force so that an increase in the short-term market interest rate will cause the money supply to increase. It will be noted also that income (Y) appears with a negative coefficient in the money supply Equation 28 which we derived in the previous section. The reason for this is that the demand for time deposits was assumed to depend on income, so that an increase in income led to an increase in time deposits, thereby reducing the amount of bank reserves available to support money balances. In the present model we neglect this effect and do not include income as a variable in the money supply function.

In this case, the following schedule shows the values of the long-term rate corresponding to the various levels of the short-term rate:

Short-term rate (Percent)	*Long-term rate (Percent)*
7.00	5.50
6.00	5.25
5.00	5.00
4.00	4.75
3.00	4.50

It is apparent that this is an approximation to the kind of relation between short-term and long-term interest rates that is depicted in Figure 14–1 or 14–3 above.

It is possible to derive *IS* and *LM* curves for the model depicted by Equations 31–39. To obtain the *IS* curve, substitute Equations 31, 32, 33, and 34 into Equation 35.

$$Y = c(Y - xY - T^*) + C_0 + i(Y - xY - T^*) - vr + I_0 + G^*$$

Solving this equation for r, we obtain the *equation of the* IS *curve*

$$r = -\frac{1 - (c + i)(1 - x)}{v}Y - \frac{c + i}{v}T^* + \frac{1}{v}(C_0 + I_0) + \frac{1}{v}G^* \qquad (40)$$

To derive the *LM* curve, substitute Equations 36 and 37 into Equation 38 to obtain

$$kY - ms + L_0 = hW^* + qs - pr_d^* + M_0 \qquad (41)$$

Solving Equation 39 explicitly for s, we obtain

$$s = \frac{1}{w}r - \frac{1 - w}{w}\bar{r} \qquad (42)$$

Substituting Equation 42 into Equation 41 and solving for r we obtain the *equation of the* LM *curve*

$$r = \frac{kw}{m + q}Y - \frac{hw}{m + q}W^* + \frac{pw}{m + q}r_d^* + (1 - w)\bar{r} + \frac{w}{m + q}(L_0 - M_0) \qquad (43)$$

To obtain the equilibrium value of Y (i.e., the value at the point where the *IS* and *LM* curves intersect), eliminate r between Equations 40 and 43 and solve the resulting equation for Y. This yields

$$Y = \frac{v}{1 - (c + i)(1 - x) + \dfrac{vkw}{m + q}}\left[\frac{1}{v}G^* + \frac{hw}{m + q}W^* - \frac{c + i}{v}T^* - \frac{pw}{m + q}r_d^* + \frac{1}{v}(C_0 + I_0) - (1 - w)\bar{r} + \frac{w(M_0 - L_0)}{m + q}\right] \qquad (44)$$

From this equation, the multipliers for government purchases and for open-market operations are easily derived

$$\frac{\Delta Y}{\Delta G^*} = \frac{1}{1 - (c + i)(1 - x) + \dfrac{vkw}{m + q}} \tag{45}$$

$$\frac{\Delta Y}{\Delta W^*} = \frac{\dfrac{vhw}{m + q}}{1 - (c + i)(1 - x) + \dfrac{vkw}{m + q}} \tag{46}$$

These multipliers are generally similar to those presented earlier (see Equations 15 and 17 in Chapter 13). However, there are two significant differences:

1. The "monetary feedback" term in the denominator of Equations 15 and 17 of Chapter 13 was "vk/m." This term has now been changed to "$vkw/(m + q)$." Since we are assuming that $0 < w < 1$ and that $m > 0$, the term "$vkw/(m + q)$" is smaller than the term "vk/m." Thus, the size of the monetary feedback is reduced by the factors considered in this chapter with the result, for example, that the multiplier for government purchases, $\Delta Y/\Delta G^*$, is increased in value. The reasons for this are easy to see. Suppose there is an increase in the level of government purchases. This will raise the level of income via the Keynesian multiplier mechanism. The rise in income will increase the demand for money, thereby tightening the money market and raising the short-term interest rate. As explained in earlier chapters, the rise in the interest rate will be smaller the more sensitive is the demand for money to the interest rate (i.e., the larger is the coefficient, m). We have now introduced another factor that will help to check the rise in the interest rate: the interest sensitivity of the *supply* of money (measured by the coefficient, q). That is, the increased pressure on the money market causes the money supply to expand (without the need for any action by the central bank) thereby helping to accommodate the increased demand for money and checking the rise in the short-term interest rate. Moreover, in the present model, a rise in the short-term interest rate caused by the tightening of the money market causes the long-term interest rate, which is assumed to affect investment spending, to rise by a smaller amount as measured by the coefficient, w.

2. The initial impact of monetary policy as shown by the numerator of Equation 17 in Chapter 13 was "v/m." This term has now been changed to "$vhw/(m + q)$" as shown in Equation 46. The introduction of the coefficient, h, is a result of the fact that in the present model the monetary policy multiplier measures the effect of a change in the central bank's portfolio of securities, W^*, rather than a change in the money stock itself, M^*, in the earlier model. The coefficient, h, reflects the ability of the banking system to engage in multiple credit expansion on the basis of an injection

of reserves and is therefore presumably greater than unity.[32] As was explained earlier, the impact of monetary policy is weakened by a large value of m, which measures the sensitivity of the demand for money to changes in the interest rate, since such sensitivity causes a change in the stock of money to have less effect on the interest rate than would otherwise be the case. In the present model, a similar effect is also produced by a large value of q, which measures the sensitivity of the money *supply* to the interest rate. And, finally, the impact of monetary policy is blunted by the fact that a change in the short-term interest rate is reflected (through the coefficient, w) in a smaller change in the long-term interest rate which is assumed to affect investment spending.

Numerical Illustration

The properties of the model we are considering can be clarified by means of a numerical illustration. Suppose the parameters of the model have the following values: $c = .8$, $C_0 = 0$, $i = .1$, $v = 15$, $I_0 = 167$, $x = .25$, $k = .2$, $m = 6$, $L_0 = 50$, $h = 4$, $q = 4$, $p = 2$, $M_0 = 10$, $\bar{r} = 5$, and $w = .25$. Thus, the model becomes

$$C = .8Y_d \quad (31a)$$
$$I = .1Y_d - 15r + 167 \quad (32a)$$
$$Y_d = Y - T \quad (33a)$$
$$T = .25Y + T^* \quad (34a)$$
$$Y = C + I + G^* \quad (35a)$$
$$M_d = .2Y - 6s + 50 \quad (36a)$$
$$M_s = 4W^* + 4s - 2r_d^* + 10 \quad (37a)$$
$$M_d = M_s \quad (38a)$$
$$r - 5 = .25(s - 5) \quad (39a)$$

Suppose the policy parameters (instruments) have the following values: $T^* = -20$, $G^* = 230$, $W^* = 40$, and $r_d^* = 5$.

Substituting the indicated values into Equation 40, we can obtain the equation for the *IS* curve

$$r = -.02167Y + 27.67 \quad (40a)$$

Likewise, substituting the parameter values into Equation 43 we obtain the equation of the *LM* curve

$$r = .005Y + 1.0 \quad (43a)$$

Solving Equations 40a and 43a simultaneously, we obtain the values, $Y = 1{,}000$ and $r = 6$. That is, the equilibrium level of income is \$1,000

[32] A comparison of Equation 37 with Equation 28 in the previous section of this chapter shows that "h" in Equation 37 corresponds to "$1/m$" in Equation 28. It will be recalled that "m" in the model of the previous section was the reserve requirement applicable to money (the weighted average of reserve requirements for demand deposits and currency).

billion and the equilibrium long-term interest rate is 6 percent. The values of all of the variables are shown in the "Original Equilibrium" column of Table 14–6.

TABLE 14–6

Effects of a $4 Billion Increase in Government Purchases
(dollar amounts in billions)

Variable	*How Calculated*	*Original Equilibrium*	*New Equilibrium*†	*Change*
Y	$C + I + G^*$	$1,000.00	$1,010.00	+$10.00
C	$.8Y_d$	616.00	622.00	+6.00
I	$.1Y_d - 15r + 167$	154.00	154.00	0.00
G^*	Exogenous	230.00	234.00	+4.00
T	$.25Y - 20$	230.00	232.50	+2.50
Y_d	$Y - T$	770.00	777.50	+7.50
M_d	$.2Y - 6s + 50$	196.00	196.80	+0.80
M_s	$4W^* + 4s - 2r_d^* + 10$	196.00	196.80	+0.80
$T - G^*$	$T - G^*$	0.00	−1.50	−1.50
W^*	Exogenous	40.00	40.00	...
s	$.02Y - .4W^* + .2r_d^* + 4$‡	9.00%	9.20%	+0.20%
r	$.25s + 3.75$§	6.00%	6.05%	+0.05%
r_d^*	Exogenous	5.00%	5.00%	...

† After increase of $4 billion in G^*.
‡ Obtained by substituting Equations 36a and 37a into 38a and solving the resulting equation for s.
§ Obtained by solving Equation 39a for r.

Now suppose there is an increase of $4 billion in government purchases of goods and services (G^*). Substituting the relevant parameter values into Equation 45, we find the value of the multiplier for government purchases

$$\frac{\Delta Y}{\Delta G^*} = 2.5$$

Thus, an increase of $4 billion in G^* will raise Y by $10 billion—from the initial equilibrium value of $1,000 billion to a new equilibrium value of $1,010 billion. The effects of the increase in government purchases on all of the variables are shown in the last two columns of Table 14–6.

In most respects, the results are similar to those obtained in the illustrations presented in Chapter 13. However, there are two new factors that are taken into account here. First, the increase in government purchases affects the maturity structure of interest rates; it raises the short-term rate by 20 basis points (.2 percentage points) while raising the long-term interest rate by only 5 basis points (.05 percentage points). Second, the increase in government purchases brings in its wake an induced increase in the money supply of $800 million. That is, even though the central bank does not change the discount rate (r_d^*), the reserve requirements of the banks (which are reflected in the value of h), or the size of its portfolio of securities (M^*), the adjustments of borrowings and excess reserves of the banks and the changes in the composition of financial assets held by the public cause the money supply to increase from \$196 billion to \$196.8 billion.[33]

Next let us consider what will happen if government purchases are held at their original level of \$230 billion while the central bank increases its portfolio of government securities (W^*) by \$4 billion to a level of \$44 billion. Substituting the relevant parameter values into Equation 46, we can obtain the value of the multiplier for open-market operations

$$\frac{\Delta Y}{\Delta W^*} = 3.75$$

Applying this multiplier to the assumed \$4 billion increase in W^*, we find that the corresponding increase in Y is \$15 billion. The effects on all of the variables are shown in the last two columns of Table 14–7.

There are two features of this illustration that should be noted. First, the open-market purchase changes the maturity structure of interest rates, lowering the short-term interest rate 130 basis points (1.3 percentage points) while lowering the long-term interest rate by only 32½ basis points (.325 percentage points). Second, the increase in the money supply resulting from

TABLE 14–7

Effects of a \$4 Billion Open-Market Purchase of Government Securities (dollar amounts in billions)

Variable	*How Calculated*	*Original Equilibrium*	*New Equilibrium*†	*Change*
Y	$C + I + G^*$	\$1,000.00	\$1,015.00	+\$15.00
C	$.8Y_d$	616.00	625.00	+ 9.00
I	$.1Y_d - 15r + 167$	154.00	160.00	+ 6.00
G^*	Exogenous	230.00	230.00	...
T	$.25Y - 20$	230.00	233.75	+ 3.75
Y_d	$Y - T$	770.00	781.25	+ 11.25
M_d	$.2Y - 6s + 50$	196.00	206.80	+ 10.80
M_s	$4W^* + 4s - 2r_d^* + 10$	196.00	206.80	+ 10.80
$T - G^*$	$T - G^*$	0.00	3.75	+ 3.75
W^*	Exogenous	40.00	44.00	+ 4.00
s	$.02Y - .4W^* + .2r_d^* + 4$‡	9.000%	7.700%	−1.300%
r	$.25s + 3.75$§	6.000%	5.765%	−0.325%
r_d	Exogenous	5.00%	5.00%	...

† After an increase of \$4 billion in W^*.
‡ Obtained by substituting Equations 36a and 37a into 38a and solving the resulting equation for s.
§ Obtained by solving Equation 39a for r.

[33] It should be noted that in this particular illustration, an increase in government purchases leaves private investment unchanged. There are two effects on investment: the increase in disposable income tends to raise it, and the rise in the long-term interest rate tends to depress it. In this particular case, these two effects exactly offset each other, leaving investment unchanged. With different values of the parameters, however, an increase in government purchases could cause investment either to rise or to fall depending on whether the income effect or the interest rate effect was the more powerful.

the open-market purchase can be broken down into two parts: (*a*) the direct effect, resulting from the coefficient of credit expansion (*h*) of 4, is to increase the money supply by $16 billion; but (*b*) the fall in the short-term interest rate of 1.3 percentage points causes adjustments by the banks and the public as reflected in the coefficient $q(=4)$ in equation (37a) which cause the money supply to fall by $5.2 billion ($1.3 \times 4 = 5.2$). Thus, the net increase in the money supply is $10.8 billion, the difference between these two effects.

FURTHER DISCUSSION OF MONETARY POLICY

The model of the economy organized around the *IS* and *LM* curves is an extraordinarily useful instrument of economic analysis. As has been demonstrated in this and the preceding chapters, a considerable number of refinements can be introduced into that model. Nevertheless, the *IS–LM* analysis has some serious shortcomings. Although we did present a rather simple dynamic version of the model and analyze its stability properties in Chapter 13, there are important dynamic elements in macroeconomics which cannot be handled within the confines of the *IS–LM* model. For example, it cannot encompass the accelerator or stock-adjustment approach to investment which was developed in Chapter 9.

The deficiencies of the *IS–LM* analysis render it an especially unsatisfactory vehicle for the discussion of monetary policy, because time lags and stock-adjustment aspects play particularly important roles in the working of monetary policy. In this section, we shall attempt to relax some of the restrictions imposed by the *IS–LM* model and discuss monetary policy in a somewhat more realistic setting. Since the complexities of a formal model are extremely forbidding, we shall have to be content with an informal treatment.[34]

Impacts of Monetary Policy

It is useful to distinguish between the initial impacts of monetary policy on aggregate demand and certain secondary and tertiary effects which occur

[34] A somewhat simplified formal model which analyzes the impact of financial variables on economic activity through the use of computer simulation is presented in W. C. Brainard and James Tobin, "Pitfalls in Financial Model Building," *American Economic Review*, Vol. LVIII, May 1968, pp. 99–122. An econometric model incorporating many of the ideas discussed in this section has been constructed by a team of economists from the Board of Governors of the Federal Reserve System and the Massachusetts Institute of Technology. Several papers have been published which describe the results of this project at various stages in its development. See Frank de Leeuw and Edward Gramlich, "The Federal Reserve–MIT Econometric Model," *Federal Reserve Bulletin*, January 1968, pp. 11–40; R. H. Rasche and H. T. Shapiro, "The F.R.B.–M.I.T. Econometric Model: Its Special Features," *American Economic Review*, Vol. LVIII, May 1968, pp. 123–49; and Albert Ando and Franco Modigliani, "Econometric Analysis of Stabilization Policies," *American Economic Review*, Vol. LIX, May 1969, pp. 296–314.

at a later stage. We begin with the initial impacts, which can be sorted out into three components.

1. ***Portfolio Adjustments.*** Monetary policy causes wealth-holders to make adjustments in their portfolios of financial and physical assets which will in due course affect income and employment. (The price level may, of course, also be affected in ways which will be discussed in some detail in Part III of this book.) The portfolio adjustment process set in motion by monetary policy may be described briefly as follows: A purchase of, say, Treasury bills by the Federal Reserve will directly lower the yield on bills and, by a process of arbitrage involving a chain of portfolio substitutions, as described earlier in this chapter,[35] will exert downward pressure on interest rates on financial assets generally. Moreover—and more important—the increase in bank reserves will enable the banking system to expand its assets. If the Federal Reserve discount rate is unchanged, the banks will be likely to use some portion of the added reserves to strengthen their free reserve position by repaying borrowings at the Federal Reserve and perhaps by adding to their excess reserves in accordance with the analysis developed earlier in this chapter.[36] But the bulk of the addition to their reserves will be used to make loan accommodation available on more favorable terms and to buy securities, thereby exerting further downward pressure on security yields.

Financial institutions other than commercial banks—such as savings and loan associations, mutual savings banks, and credit unions—are likely to become involved in the credit expansion process as some portion of the funds generated by credit expansion are deposited in these institutions. The proportions in which the expansion of credit is reflected in increases in the public's holdings of currency, demand deposits, time deposits, savings and loan shares, and such, will be determined by the movements in interest rates on securities and on the various types of interest-bearing deposit claims and the extent to which the public's demands for these various types of claims respond to changes in these interest rates.[37]

With the expected yield on a unit of real capital—that is, the marginal product of capital corresponding to the existing stock of capital—unchanged, the decline in the yields on financial assets and the more favorable terms on which new debt can be issued will throw the balance sheets of households and businesses out of equilibrium. The adjustment toward a new portfolio equilibrium will take the form of sales of financial assets and the issuance of new debt to acquire real capital and claims thereto. This will raise the price of existing units of real capital—or equity claims against these units—

---

[35] See pp. 280–94 above.

[36] See pp. 298–99 above.

[37] A somewhat simplified version of this determination was discussed earlier in this chapter in the discussion of a model involving money, time deposits, and securities. See pp. 295–309 above.

relative to the initially unchanged cost of producing new units, thereby opening up a gap between desired and actual stocks of capital, a gap that will gradually be closed by the production of new capital goods.

This stock-adjustment process was discussed in some detail in Chapter 9 in its application to business investment in plant and equipment. However, it can also be applied, with some variations to suit the circumstances, to a variety of other forms of investment, including business investment in inventories and household investment in new homes and in consumer durable goods such as automobiles, electrical appliances, and furniture. There is evidence that all of these components of aggregate demand, as well as expenditures on schools, highways, and other public facilities by state and local governments, are affected in varying degrees by monetary policy.[38]

2. ***Wealth Effects.*** Since monetary policy operates entirely through voluntary transactions involving swaps of one financial asset for another, it does not add to wealth by creating assets to which there are no corresponding liabilities. Nevertheless, monetary policy does have wealth effects, which may be of considerable importance. As was explained in Chapter 11, the yields on existing bonds are kept in line with changing market interest rates through changes in the prices of those bonds. Thus a fall in interest rates will increase the market value of wealth by causing bond prices to rise. Similarly, the market values of a variety of other income-yielding assets are determined by capitalizing the income streams expected from these assets at appropriate rates. A monetary policy which reduces interest rates generally will lower capitalization rates, thereby causing a rise in the market value of such assets as real estate and common stock equities.[39] In part, this may strengthen the impact on economic activity of the portfolio adjustments already referred to by increasing the size of the portfolios available for allocation. For example, an investor experiencing a capital gain through appreciation of the value of bonds may realize the capital gain by selling the bonds and invest the proceeds in real capital. In addition, the increase in household wealth may stimulate consumption. It was shown in Chapter 5 that, according to the so-called life cycle hypothesis, consumption spending may be affected by both disposable personal income and household net worth

[38] For a useful survey of recent empirical work relating to the effects of monetary factors on various components of demand, see M. J. Hamburger, "The Impact of Monetary Variables: A Survey of Recent Econometric Literature," in *Essays in Domestic and International Finance* (New York: Federal Reserve Bank of New York, 1969), pp. 37–49, reprinted in Smith and Teigen, *Readings* (rev. ed.) *op. cit.*, pp. 414–33. Many of the studies referred to in this paper involve the use of some form of stock-adjustment model.

[39] Consider a piece of property which is expected to yield a rent, after expenses, of \$1,000 a year for an indefinitely long period into the future. If the rate of return that can be earned on other investments judged to involve equivalent risk is 8 percent, the market value of this property will be \$12,500 (1,000 ÷ .08 = 12,500). If a general decline in interest rates causes the relevant capitalization rate to fall to 7 percent, the market value will rise to \$14,286 (1,000 ÷ .07 = 14,286).

or wealth. There is some evidence that the wealth effect on consumer spending brought about by a decline in interest rates may be of considerable importance.[40]

3. ***Credit Availability Effects.*** The portfolio and wealth effects are the basic channels through which monetary policy has its initial impact on economic activity. In addition, however, the institutional arrangements for providing funds to certain sectors may give monetary policy a special leverage over the *availability* of credit to these sectors, thereby affecting their ability to spend. Any attempt to discuss credit availability effects in detail would involve an analysis of the financial sector of the U.S. economy which is beyond the scope of this book. However, the sector which seems most clearly to be affected substantially by credit availability is residential construction, and it will be useful to sketch briefly the way in which this sector is affected. It is perhaps most illuminating to discuss the way in which a restrictive monetary policy may curtail the availability of funds for homebuilding, although an expansionary policy will, of course, have the same effects in reverse.

Even in the absence of the rather unique institutional arrangements for its financing, housing demand might be significantly affected by monetary policy as changes in mortgage interest rates altered the desired housing stock. But as postwar experience has repeatedly shown, changes in mortgage credit availability may greatly strengthen the impact of restrictive monetary policy on homebuilding and cause the effects to occur much more rapidly than the stock-adjustment mechanism would imply. There are two important ways in which mortgage credit availability may be affected by a restrictive monetary policy.[41]

First, when the Federal Reserve adopts a restrictive policy to curtail excessive aggregate demand, the commercial banks may raise interest rates on time and savings deposits in order to attract funds to satisfy the needs of their business customers. If savings and loan associations do not raise the rates paid to their depositors or raise them less than the banks raise their rates, households may rechannel their savings flows away from savings and loan associations and toward banks—or may even withdraw existing savings from savings and loan associations and shift them to banks. The Federal Reserve may try to prevent such a shift of funds from occurring by setting

---

[40] In some recent simulations performed with the FRB–MIT econometric model (see footnote 34 above), the effect on consumption resulting from the induced change in the value of common stock equities held by households accounts for 35 to 45 percent of the initial impact of monetary policy.

[41] For a more extensive discussion, see the paper by the staff of the Federal Home Loan Bank Board, "Cycles in Mortgage Credit Availability and the 1966 Experience," in *A Study of Mortgage Credit,* Subcommittee on Housing and Urban Affairs, Committee on Banking and Currency, United States Senate (Washington: U.S. Government Printing Office, 1967), pp. 19–40, reprinted with minor changes in Smith and Teigen, *Readings* (rev. ed., *op. cit.*), pp. 433–47.

ceilings (as it is legally empowered to do) on the rates the banks may pay on time and savings deposits in order to shield the savings and loan associations from bank competition. Even if this is done, however, a rise in interest rates on marketable securities may cause households to shift funds away from fixed-value redeemable claims generally—including deposits in both banks and savings and loan associations—and directly into the securities markets. Savings and loan associations are highly specialized institutions which—as a result of both legal restrictions and customs—channel almost all of their deposit inflows into the mortgage market. To the extent that they lose funds either to commercial banks or to the securities markets, the availability of mortgage credit will be directly curtailed, with a powerful restrictive effect on homebuilding activity. It might appear that if savings and loan associations experienced intensified competition for funds, they could simply raise the interest rates paid to their depositors in order to attract or hold the funds needed to maintain their mortgage lending activity. However, in a period of sharply rising interest rates, they are likely to find themselves with frozen portfolios of older mortgages which pay lower interest rates than currently prevail, making it difficult for them to pay markedly increased interest rates to their depositors.

Second, as interest rates rise, yields on corporate bonds typically rise relative to interest rates on new mortgages. As a consequence, some institutional investors, such as life insurance companies, shift the composition of their investment flows away from mortgages and toward corporate bonds. This tendency may be exacerbated by legal ceilings that are sometimes applied to mortgage interest rates. Such ceilings have existed, for example, on mortgages insured or guaranteed by the Federal government under programs operated by the Federal Housing Administration (FHA) and the Veterans Administration (VA). In addition many states have usury laws which impose unrealistically low ceilings on the interest rates than can be charged for mortgage loans.

Other sectors besides homebuilding may be affected to some degree by changes in credit availability. Many state and local government jurisdictions have statutes or administrative regulations which limit the interest rates that may be paid on securities issued to finance schools, highways, or other public facilities. Such limitations may, on occasion, prevent jurisdictions from issuing bonds to finance such projects even though their citizens would be willing to pay prevailing interest rates for the money.[42]

[42] For a discussion of the effects of restrictive monetary policy in 1966 on borrowing and capital spending by state and local government units, see P. F. McGouldrick and J. E. Petersen, "Monetary Restraint and Borrowing and Capital Spending by Large State and Local Governments in 1966," *Federal Reserve Bulletin,* July 1968, pp. 552–81; and J. E. Petersen and P. F. McGouldrick, "Monetary Restraint, Borrowing, and Capital Spending by Small Local Governments and State Colleges in 1966," *Federal Reserve Bulletin,* December 1968, pp. 953–82. A summary of these two papers is reprinted under the title, "Effects of Monetary Restraint on State and Local Governments in 1966," in Smith and Teigen, *Readings* (rev. ed., *op. cit.*), pp. 449–75.

Secondary Effects: The Multiplier and Accelerator

The initial impacts of an expansionary monetary policy working through the portfolio effects, wealth effects, and credit availability effects discussed above will increase spending and generate income. The associated rise in disposable income will further increase the demand for consumer nondurable goods and services through the operation of the multiplier as discussed in Chapter 6. It will also increase the demand for the services of houses and consumer durable goods. The increased demand for both consumer and capital goods will also raise the stock of capital desired by business, thereby setting in motion the accelerator or stock-adjustment process described in Chapter 9. Thus, the familiar magnification of demand through multiplier and accelerator effects will be brought into play.

Tertiary Effects: The Monetary Feedback

A tertiary chain of effects is set in motion as the rise in income generated by the initial and secondary effects increases the demand for money for transactions purposes. This will reverse the initial decline in interest rates, and the induced rise in interest rates will exert a dampening effect on the expansion by a partial reversal of the portfolio effects, wealth effects, and credit availability effects which initially triggered the rise in income.

Concluding Comments

It is not possible in practice to draw a clear and sharp distinction between the initial, secondary, and tertiary effects discussed above. This is because all of the effects occur with time lags which cause them to build up gradually over time. Since the first-stage initial impacts set in motion secondary and tertiary effects which are occuring at the same time as the later stages of the initial impacts, the whole process in practice becomes inextricably interwoven. Moreover, since the lags at each stage may be quite long and complex, the process may be very slow to work itself out.[43] The fact that the lags appear to be very long—and may also be subject to erratic variation from one situation to another—has raised a question in the minds of some economists about the efficacy of discretionary monetary policy aimed at economic stabilization, suggesting that a policy aimed to produce a steady growth of the money supply or bank reserves might be more suitable. While the present author believes the use of discretionary policy is preferable to

---

[43] The fact that the lags involved in the working of monetary policy are quite long is clearly indicated in many of the simulations performed with the FRB–MIT econometric model. See the literature relating to this model cited in footnote 34 above.

such a so-called monetary rule, a detailed discussion of this controversy is beyond the scope of this book.[44]

It should finally be noted that all three of the effects of monetary policy —the initial impact, the secondary effect, and the tertiary effect—are reflected in a rudimentary way in the *IS–LM* model developed in this book. Consider, for example, the multiplier for open-market operations which was presented earlier in this chapter (Equation 46). For convenience, this multiplier is repeated.

$$\frac{\Delta Y}{\Delta W^*} = \frac{\dfrac{vhw}{m+q}}{1-(c+i)(1-x)+\dfrac{vkw}{m+q}}. \tag{46}$$

The numerator of this expression, $vhw/(m+q)$, represents the initial impact of monetary policy; the terms at the left-hand end of the denominator, $1-(c+i)(1-x)$, represent the secondary—i.e., multiplier—effects; and the last term in the denominator, $vkw/(m+q)$, represents the tertiary —i.e., monetary feedback—effect.

---

[44] The author has expressed his views on a number of controversial issues relating to monetary policy, including this one, in "A Neo-Keynesian View of Monetary Policy," in *Controlling Monetary Aggregates* (Boston: Federal Reserve Bank of Boston, 1969), pp. 105–26. For a classic statement of the case in favor of monetary rule, see Milton Friedman, "The Role of Monetary Policy," *American Economic Review,* Vol. LVIII, March 1968, pp. 1–17. Arguments on the other side are presented in L. E. Gramley, "Guidelines for Monetary Policy—The Case Against Simple Rules," paper delivered at the Financial Conference of the National Industrial Conference Board, New York, February 21, 1969. These two papers are reprinted in Smith and Teigen, *Readings* (rev. ed., *op. cit.*), pp. 476–88 and 488–95, respectively.

PART III

Prices, Wages, and Employment

In our discussion of aggregate demand and income determination in Part II, we deliberately avoided dealing with the price level by assuming that the economy was operating at substantially less than full employment with a large amount of excess productive capacity, and that under these conditions an increase in aggregate demand would cause a rise in production and employment without an appreciable rise in wages and prices. Although such an analysis might not be seriously inappropriate in dealing with an economic system in the midst of a serious depression such as that of the 1930's, it is obviously quite unsatisfactory as applied to the postwar period, when the American economy has much of the time operated reasonably close to capacity with relatively little unemployment and when we have at times been forced to curtail aggregate demand in order to prevent inflation.

Before the appearance of Keynes' *General Theory,* the dominant view among economists was that a capitalist economic system contained a powerful mechanism which generally tended to keep it operating at or near full employment without the need for governmental stabilization policies. In Chapter 15, we consider the way in which this mechanism was supposed to work and then examine the defects that make it ineffective in the modern world.

Chapter 16 deals with inflation. After considering briefly the economic effects of inflation, we take up inflation that results simply from excessive aggregate demand and then turn to a more realistic discussion of the way in which the existence of market imperfections and market power by large business and labor organization can complicate and exacerbate the problem of inflation.

Chapter 15 PRICE FLEXIBILITY AND EMPLOYMENT

The way in which a predominantly capitalist economic system such as ours functions, and particularly the way in which it responds to changes in the money supply, depends to a considerable extent upon the flexibility with which the wages of labor and the prices of goods respond to changes in conditions of demand and supply prevailing in the various markets. Let us now consider some of the possibilities in this respect.[1]

FLEXIBLE PRICES AND WAGES AND AUTOMATIC FULL EMPLOYMENT

We must begin our analysis by considering the operation of the labor market. In a free enterprise economy, the demand for labor is based on the marginal productivity principle. Under competitive conditions, the profit maximizing producer will carry employment to the point where the money wage is equal to the value of the marginal product of labor. The value of the marginal product is equal to the marginal physical product of labor—i.e., the addition to total output of the product resulting from the employment of the last worker—multiplied by the price of the product. Thus, the value of the marginal product is the addition to the firm's revenue resulting from the employment of the last worker. Under competitive conditions, the firm will maximize its profits if it hires all workers for whom the value of the marginal product is greater than the money wage rate—that is, all workers who add more to the firm's revenue than to its cost.[2] The

[1] For a fuller discussion of these matters, see W. L. Smith, "A Graphical Exposition of the Complete Keynesian System," *Southern Economic Journal,* Vol. XXIII, October 1956, pp. 115–25, reprinted in W. L. Smith and R. L. Teigen (eds.), *Readings in Money, National Income, and Stabilization Policy* 2d ed.; (Homewood, Ill.: Richard D. Irwin, Inc., 1970).

[2] If the firm enjoys a monopolistic position in selling its product, so that it is faced with a noticeably downward sloping demand curve, the marginal revenue from the sale of its product will lie below the price of the product, and it will carry employment to the point where the money wage is equal to the marginal revenue product of labor, which is obtained by multiplying physical product by the marginal revenue. (In the competitive case, the price of the product is equal to the marginal revenue, so that the value of the

marginal physical product of labor declines as employment increases, since, with the factors of production other than labor held constant, each unit of labor has a smaller amount of other factors to assist it as the amount of labor employed is increased. If w is the money wage rate, p is the price of the product, and *MPP* is the marginal physical product, we have

$$w = MPP \times p$$

or

$$\frac{w}{p} = MPP$$

Since the *MPP* declines as employment increases, it follows that the amount of employment increases as the "real wage" (w/p) declines. Extending this principle to the economy, we can draw a downward-sloping demand curve for labor, as depicted by the curve *DD* in Panel A in the lower lefthand section of Figure 15–1. This curve shows for each "real wage," measured vertically, the amount of employment, measured horizontally, that will be offered by the employers of the economy.

The supply of labor is also assumed to depend upon the real wage and is depicted by the curve *SS* in Panel A of Figure 15–1. This curve indicates that the amount of employment that workers will be willing to supply will increase as the real wage rises. It is drawn quite steeply in order to depict the fact that working hours are quite commonly fixed by convention in our economy and cannot readily be varied at the option of the individual worker, as well as the fact that at any given time the number of workers who will be willing to work is probably determined mainly by the number of family units to be supported and does not vary a great deal as the real wage changes over the usual range of variation. However, the exact shape of the supply curve of labor is not very important for our present purposes.

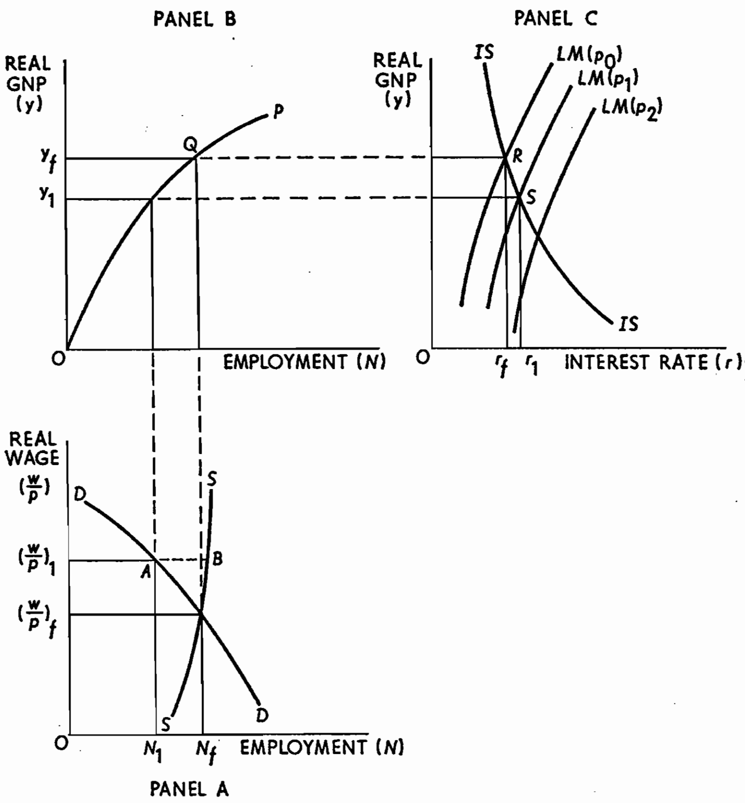
The interaction of the supply and demand for labor as depicted in Panel A of Figure 15–1 determines the amount of employment (designated by N) and the real wage (designated by w/p). Thus, employment will be N_f and the real wage will be $(w/p)_f$. This can be described as a situation of "full employment" (this is the meaning of the subscript "f" attached to N and w/p), since when employment is N_f supply and demand for labor are equated. Then there are no workers willing to work at the prevailing

---

marginal product is equal to the marginal revenue product.) If the firm is also in a monopsonistic position, so the money wage rises visibly as the firm increases its employment, the marginal cost of labor will exceed the wage and the firm will carry employment to the point where the marginal cost of labor equals the marginal revenue product. For a more extensive discussion of these matters, see G. F. Bloom and H. R. Northrup, *Economics of Labor Relations* (4th ed.; Homewood, Ill.: Richard D. Irwin, Inc., 1961), chap. 11. It is possible to adjust the analysis presented in Figure 15–1 to allow for the existence of monopoly, monopsony, or both.

FIGURE 15–1

Wage and Price Adjustments in Relation to Income and Employment



real wage who do not have jobs and at the same time employers have all the labor they are willing to hire at the prevailing real wage.

Since we are assuming that the stock of capital equipment, the supplies of natural resources, and the state of technology are all given, the only variable factor of production is labor; accordingly there is a specific level of real GNP associated with each level of employment.[3] The relation between employment and real GNP is indicated by the curve *OP* in Panel B of Figure

[3] It should be noted that our analysis treats the economy as though it produced only one good. Thus, we assume a precise relation between employment and output. In reality, the same level of employment can produce many different combinations of goods and services depending on how it is allocated among firms and industries. This means that there is a problem of measuring aggregate output, which was discussed briefly in Chapter 4, when we dealt with the construction of index numbers and the problem of deflation for price changes. It is felt, however, that the behavior of a multiproduct economy bears a close enough similarity to the behavior of a hypothetical economy producing one product to render the present analysis useful.

15–1. The tendency of this curve to flatten out as employment increases reflects the fact that GNP increases less than in proportion to increases in employment due to the operation of the law of diminishing returns.[4]

The level of employment having been determined by the supply and demand for labor as shown in Panel A, it is possible to find the level of real GNP by tracing up vertically from N_f in Panel A to the *OP* curve at point *Q* in Panel B and then tracing horizontally from *Q* to the vertical axis of Panel B. Thus, the level of real GNP is y_f (the subscript "*f*" designating the "full employment" level of real GNP).

The *IS* and *LM* curves for the economy are depicted in Panel C of Figure 15–1. The student should note, however, that the diagram has been turned on its side, with the interest rate measured on the horizontal axis and income (real GNP) on the vertical axis. The reason for this change is, of course, to make the diagram "mesh" with Panels A and B of Figure 15–1.

The *IS* curve in Panel C shows the various combinations of real GNP and interest rate at which the market for goods and services will be in equilibrium, i.e., at which aggregate demand for goods and services will be equal to aggregate supply. Thus, tracing across from Panel B to Panel C at the GNP level y_f, we intersect the *IS* curve at point R. Tracing down from this point to the horizontal axis, we find that the interest rate must be r_f (the subscript "*f*" designating the "full employment" interest rate) if the market for goods and services is to be in equilibrium at income y_f.

In our discussion of the interest rate in Chapter 11, we assumed that the price level did not change, so it was not necessary to deal with the price level in relation to the interest rate. In our present analysis, however, the price level is a variable. As the price and wage levels decline, the demand for money for transactions purposes will decline—i.e., it will take smaller cash balances to meet payrolls at low than at high wages and to meet grocery bills at low than at high prices. Thus, the demand for money is presumably a demand for a certain amount of purchasing power rather than for a certain number of dollars.

We may designate the purchasing power represented by the money stock *M* as M/p—i.e., the real value of the money stock. Now this amount of purchasing power may be increased by increasing the nominal amount of money *M* or by a fall in the price level *p* which will increase the purchasing power of the given nominal money stock. As was shown in Chapters 12 and 13, an increase in the quantity of money will shift the *LM* curve upward, while a reduction in the money supply will shift the *LM* curve downward. (It should be noted that changes in the supply of money will shift the *LM* curve to the left or right in the diagrams in Chapter 13, such as Figure 13–3, whereas they shift the curve up or down in Figure 15–1; this is because the axes are reversed as between the two diagrams.) Thus, with a given supply

[4] The mathematically inclined student will note that the curve *DD* in Panel A is the derivative of the curve *OP* in Panel B.

of money, there is a different *LM* curve for each price level; the higher the price level the lower is the *LM* curve in Panel C of Figure 15–1. For purposes of illustration, three such curves are shown—the curves labeled $LM(p_2)$, corresponding to price level p_2; $LM(p_1)$, corresponding to price level p_1; and $LM\ (p_0)$, corresponding to price level p_0. Of course, price level p_2 is higher than p_1, and p_1 is higher than p_0—that is, the curve shifts upward on the chart as the price level falls. In the example shown, the price level p_0 is the full employment equilibrium price level, since the *LM* curve corresponding to this price level intersects the *IS* curve at point *R*, which corresponds to full employment. With flexible prices and wages, the price level will tend to settle at p_0, thus establishing full employment.

To illustrate the mechanism by which the economy tends automatically to establish full employment, consider what would happen if for some reason the economy should be operating below full employment. Suppose, for example, that the real wage were $(w/p)_1$ (higher than the full employment real wage $(w/p)_f$) and employment were N_1 (less than the full employment level N_f) in Panel A. Accordingly, suppose the GNP were y_1 (below the full employment level y_f) in Panel B and that the price level were p_1 (above the full employment level p_0) so that the *LM* curve took the position $LM(p_1)$ in Panel C. As long as wages and prices and wages are flexible—i.e., as long as wages fall when the supply of labor exceeds the demand and the prices of goods fall when the supply of goods exceeds demand—the situation just depicted cannot be maintained. Looking at Panel A, the supply of labor exceeds demand by the amount *AB* at the real wage of $(w/p)_1$; that is, there are unemployed workers who would like to work at the prevailing (or even a lower) wage. In an effort to obtain employment, these workers will bid down the money wage. As the money wage falls, costs of production of goods and services will decline, and if the prices of goods and services are flexible they will fall. As prices and wages fall, the amount of money needed for transactions to handle payrolls, grocery bills, and so on, is reduced. Thus, the supply of money comes to exceed the demand at the prevailing interest rate. The excess money is put into the loan market, by the purchase of securities or by paying back bank loans, thus leaving the banks with unused lending capacity and inducing them to offer loans on better terms. As a result of this, interest rates fall, and the fall in interest rates stimulates investment to the extent that the marginal efficiency of investment is interest-elastic. The increase in investment raises aggregate demand, both directly and through the multiplier mechanism. The increase in demand will induce an increase in production and employment to meet it. This process will continue as long as the supply of labor exceeds the demand for it, since as long as this condition prevails, money wages will continue to fall. The process will come to a stop, when employment reaches N_f, the real wage reaches $(w/p)_f$, real GNP reaches y_f, the interest rate reaches r_f, and the price level reaches p_0.

THE REAL BALANCE OR "NET CLAIMS" EFFECT

In addition to the effects discussed above, some monetary theorists have emphasized another channel through which flexible wages and prices may work to maintain full employment in the economy. These theorists contend that as the real value of wealth increases, the desire to save to accumulate further wealth weakens and households (and perhaps firms as well) tend to spend more and save less at given levels of income. That is, an increase in consumer wealth tends to raise expenditures, thus exerting a stimulating effect on the economy.[5]

Now the market values of real goods and of equities in enterprises producing goods and services tend in general to move roughly in proportion with the price level of currently produced goods and services; consequently, a decline in prices has little effect on the aggregate value of real wealth. However, debt claims—which constitute promises to pay fixed sums in money—as well as money itself, increase in real value as the price level falls; that is, their real value in terms of currently produced goods and services increases. Thus, it might be argued that a fall in the price level, by increasing the real value of that portion of their wealth which people hold in the form of debt claims and money, would cause a rise in expenditures.

The difficulty with this argument, however, is that private debt instruments are not only an asset to the person holding them but a liability to the person or firm issuing them. A fall in the price level thus not only increases the real value of the asset to the person holding it but at the same time increases the real burden of the debt to the person owing it. If the interest and principal of the debt are not paid when due, the person owing the debt may be forced into bankruptcy with disastrous results for him. Consequently, a fall in the price level will make debtors more cautious and have a depressing effect on their expenditures which is likely to cancel out—and perhaps more than cancel out—the stimulating effect on expenditures of the rise in the real value of the corresponding assets to their owners. This same argument applies to a large part of the money stock as well as to debt claims, since much of the money in the economy is created by the making of loans to private spending units. In such cases, the money holdings of one person are balanced by debt owed to the banking system by someone. Thus, a fall in prices may encourage the holders of money to spend more while at the same time discouraging spending by those owing debt to the banks.

There is, however, a stock of monetary and debt claims held by spending units in the economy which does not have a counterpart in debts owed by other spending units. Such claims arise when the federal government spends

[5] There is a very extensive literature on this subject, the most comprehensive treatment (though at a quite advanced level) being contained in Don Patinkin, *Money, Interest, and Prices* (2d ed.; New York: Harper & Row, 1965).

more on goods and services or transfer payments than it takes in in taxes revenues, i.e., whenever the government has a budget deficit.[6] It should be noted that such "net claims against the government" are generated by federal deficits no matter how the deficits are financed. If the government raises money by selling bonds to the public, it gives the public bonds in exchange for money; when it spends the money, the public's holdings of money are restored to the previous level. Thus, when the whole operation of borrowing and spending is completed, the public has the same amount of money as it had to begin with and an additional amount of claims in the form of government bonds in an amount equal to the deficit. If bonds are sold by the government to the banks, who create deposits to the credit of the government to buy them, and the government then spends the deposits so created on goods and services, the public will have more money (i.e., deposits) without any increase in its indebtedness to the banks—that is, the public's net claims are increased. Finally, if the government finances the deficit by printing money or by drawing down its previously accumulated cash balances, the public's holdings of money are increased without any increase in its indebtedness. Thus, *federal budget deficits increase the public's net claims dollar for dollar, and, of course, federal surpluses (excesses of tax receipts over expenditures) reduce net claims dollar for dollar.*

Thus, the nominal value of the stock of net claims in existence at any time is equal to the cumulative net budget deficit of the federal government since the beginning of its operations. The amount of purchasing power represented by this stock of net claims increases as the price level falls and falls as the price level rises. If changes in the price level cause changes in the public's propensity to spend on goods and services, it is presumably a result of such changes in the purchasing power to the public's net claims.

This alleged tendency for price fluctuations to produce inverse changes in expenditures as a result of changes in the public's wealth has been variously called the "cash balance effect," the "real balance effect," and the "Pigou effect" (after the late British economist A. C. Pigou, who first called attention to the possibility). The first two names are somewhat misleading since they suggest that the effect applies to money balances, which for reasons indicated above is hardly correct. Perhaps the best name that could be applied to the phenomenon would be "net claims effect" in accordance with the reasoning set forth above.

[6] Changes in net claims result from deficits and surpluses of the federal government but not from those of state and local governments. The federal government has the ultimate power to create money if necessary to pay off its debts and therefore cannot be rendered bankrupt. Consequently, there is no reason why a fall in prices which increases the real burden of the federal government debt should create pressure on that government to curtail its expenditures. State and local governments, on the other hand, do not have the power to create money, and their securities are accordingly subject to some risk of default, like those of private spending units. A fall in prices therefore increases the debt burden on state and local governments and may create pressure for them to curtail expenditures, just as it does for private debtors.

However, even as applied to the public's net claims, the operation of the alleged effect encounters some logical difficulties, and its importance has never been clearly demonstrated empirically. The question of price expectations is one source of trouble. When prices are declining during a period of depression, it seems probable that the public will regard these price declines as merely temporary phenomena which are likely to be reversed when recovery occurs. If this is the case, it seems most unlikely that people will revise their calculations concerning the real value of their wealth on the basis of such price declines—that is, it does not seem reasonable that the public will feel richer under such circumstances and therefore be inclined to increase expenditures on goods and services. Moreover, the problems of offsetting financial claims and debts generated by private borrowing and lending are not satisfactorily disposed of by simply canceling the totals out against each other and taking the net claims as suggested above. The fall in prices makes creditors feel better off and debtors feel worse off, but there is no reason to assume that these effects cancel out. In fact, in the absence of any very firm evidence, one might hazard a guess that the depressing effect on debtors, who are under severe compulsion to meet the interest and principal of their debts, will often be more powerful than the buoyant effects on creditors, who are under no strong pressure to increase their spending.

It is important to note that the stock of net claims held by the public may be increased by government deficits which result in increases in the nominal dollar amount of claims, as well as by declines in the price level which increase the real value of a fixed nominal amount of claims. A rise in the nominal amount of claims may have a much greater effect on spending than an increase in the real value of claims resulting from a fall in prices. During the World War II period of June 30, 1939, to June 30, 1946, the federal government operated at a cumulative cash deficit of $190 billion, thus increasing the public's net claims by this amount. There is little doubt that the high levels of consumer and business demands that prevailed in the immediate postwar period were partly due to this vast accumulation of financial assets by the public as a byproduct of war finance. Under normal peacetime circumstances, on the other hand, it seems doubtful whether the relatively small budget deficits and surpluses that usually occur are capable of changing the public's holdings of claims by sufficiently large amounts to be very important.

Such evidence as there is suggests that the "net claims effect" is not ordinarily very important, and we shall not be inclined to take it seriously. To the extent that the effect is present, a fall in prices will tend to cause a rise in the consumption function and perhaps also in investment spending. Thus, with a net claims effect present, a fall in prices would cause the *IS* curve in Panel C of Figure 15–1 to shift upward and a rise in prices would cause the *IS* curve to shift downward. If the economy were temporarily at point *S* in Panel C with substantial unemployment, as in our earlier example, the resulting decline in wages and prices would now work through two different channels to restore full employment: (1) the rise in the real value of cash balances would cause the *LM* curve to shift upward, for reasons explained earlier, and (2) the increase in the real value of net claims would cause the *IS* curve to shift upward.

TWO CAUSES OF CHRONIC UNEMPLOYMENT

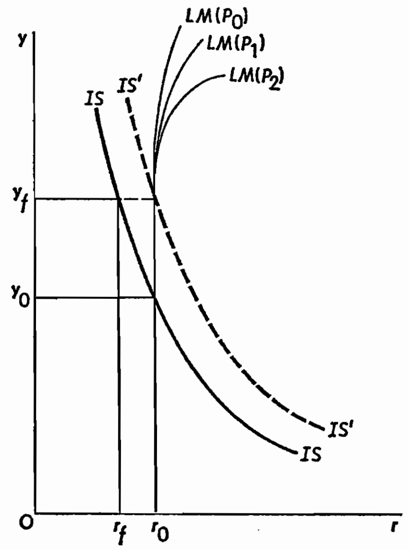
The above analysis is intended to show that a free enterprise economy has built into it an automatic tendency to maintain full employment or to restore itself to full employment if a situation of unemployment should prevail temporarily. However, there are two important ways in which the forces of equilibration may be at least temporarily immobilized.

The "Liquidity Trap" Problem

It was pointed out in Chapter 11 that if the speculative motive for holding money is operative and important, there may be a floor below which it will be extremely difficult to force the rate of interest. The reason is that if investors are influenced by conventional views concerning the "normal" level of interest rates, it may be that after the rate has fallen substantially below this level, investors will become virtually unanimous in their feeling that a rise in the rate is in the offing, and they will be unwilling to invest further resources in bonds no matter how much the supply of money is increased. If this should occur, a further decline in prices which increases the real value of cash balances will be completely ineffective in lowering the rate of interest and thereby stimulating investment.

Such a situation is depicted in Figure 15–2, which is analogous to Panel C of Figure 15–1. The full employment level of real GNP as determined by the supply and demand for labor and by the production function (Panels A and B of Figure 15–1) is y_f. The interest rate necessary to make aggregate demand equal to y_f, as indicated by the *IS* curve in Figure 15–2, is r_f. However, the liquidity preference schedule becomes completely elastic at an interest rate of r_0, which is higher than r_f. When the liquidity preference schedule becomes perfectly elastic, the *LM* curve, as drawn in Figure 15–2, becomes a perfectly vertical line at the same interest rate. (The student should convince himself of this by studying the derivation of the *LM* curve from the liquidity preference schedule as shown in Figure 12–6.) The *LM* curves for various price levels under these conditions are shown in Figure 15–2. A fall in the price level raises the *LM* curve, but does not change the location of the vertical section of it. Since, in this case, the *IS*

FIGURE 15–2
The "Liquidity Trap" Case



curve intersects the *LM* curve in the vertical range of the latter, a fall in the price level does not reduce the interest rate as it did in Figure 15–1, and therefore does nothing to raise output and employment.

It may be noted that if there is a significant "net claims effect" present, this effect may produce a tendency toward full employment even though the interest rate effect is immobilized. That is, due to the "net claims effect" a fall in the price level may cause the *IS* curve to shift upward; if this process continues until the curve occupies the position *IS′*, full employment will be achieved. However, for reasons suggested earlier, the importance of the "net claims effect" is rather dubious.

Rigidity of Money Wages

Another way in which the automatic forces tending to produce full employment may be thwarted is through the existence of downward rigidity of money wage rates. Such rigidity of money wages may be due to various factors, including (*a*) powerful trade unions which are able to prevent money wages from falling, at least temporarily; (*b*) statutory provisions, such as minimum wage laws; (*c*) failure of employers to reduce wages even though besieged by workers willing to work at lower wages, due to a desire to retain loyal and experienced employees and to maintain morale; or (*d*) unwillingness of unemployed workers to accept reduced money wages even though they would be willing to work at lower real wages brought about by a rise in prices.[7]

If money wages are rigid, it is possible for the economy to get into a position that can be described as an "underemployment equilibrium." Such a situation would look like that shown in Figure 15–1 in which employment is N_1, the real wage is $(w/p)_1$, real GNP is y_1, and the price level is p_1, so that the *LM* curve occupies the position $LM(p_1)$. If the fixed money wage (w) is such that when it is deflated by the price level p_1 the real wage is at the level $(w/p)_1$, this will represent an equilibrium position that will be maintained. (As indicated in our earlier discussion, if wages are flexible so that they will fall due to the existence of the unemployment—which amounts to *AB* in Panel A of Figure 15–1—the position will not be an equilibrium, because the resulting fall in money wages will set in motion forces that will lead to full employment.)

In principle at least, unemployment that is due to wage rigidity can be cured by an appropriate increase in the money supply. In the example shown in Figure 15–1, all that is necessary is to increase the money supply sufficiently to shift the *LM* curve so that it intersects the *IS* curve at point *R* when the price level is high enough to produce full employment. This price level consistent with full employment is one which is high enough to convert the fixed money wage into a real wage of $(w/p)_f$, the level of real wages consistent with equilibrium in the labor market.

THE PRACTICAL SIGNIFICANCE OF ALL THIS

Prior to the Great Depression of the 1930's (and especially prior to the appearance of Keynes' *General Theory of Employment, Interest, and Money* in 1935), it was customary for economists to emphasize that a free enterprise economy possessed strong built-in automatic tendencies toward full employment. Large-scale unemployment was usually regarded as a distinctly unimportant temporary aberration of the system, and very little attention was given to it in the writings of most professional economists. Today, however, there is much greater awareness of the complexity of the automatic self-corrective forces tending to prevent unemployment or to eliminate it once it has developed.

In practice, it is quite unlikely that either of the two basic flaws in the equilibration mechanism referred to above will be present in extreme form. That is, a true "liquidity trap" situation in which the demand for money becomes *completely* interest-elastic seems rather improbable. It is possible

[7] It may appear irrational for workers to be unwilling to accept a cut in money wages but to be willing to accept a cut in real wages brought about by a rise in prices. However, if workers have commitments that are fixed in terms of money—such as mortgage payments they have to make—a distinction of this kind may not be irrational. On this general subject, see James Tobin, "Money Wage Rates and Employment," in S. E. Harris (ed.), *The New Economics* (New York: Alfred A. Knopf, 1947), pp. 572–87.

that such a situation did occur in the United States in the 1930's, but even this is debatable. Nor are money wages likely to be *completely* rigid; when unemployment becomes large, some decline in money wages is almost certain to occur.

Nevertheless, both of the difficulties referred to above appear to be present in somewhat attenuated form. Due to trade union pressures, minimum wage laws, and so on, money wages do not decline readily when unemployment occurs; while not completely inflexible, wages may be described as "sticky" rather than truly flexible in a downward direction. That is, moderate amounts of unemployment seem capable of causing no more than a slow and gradual sag in money wages. Now, as indicated earlier, a decline in wages (and prices) raises the level of aggregate employment in the economy by increasing the real value of the money stock, lowering the interest rate, and perhaps by increasing the real value of the stock of "net claims against the government," thereby causing an upward shift in expenditure schedules.

While the demand for money is most unlikely to be completely interest elastic, it does appear that the interest-elasticity of demand for money may commonly be great enough so that a rather large increase in the real value of cash balances—and therefore a large decline in money wages—will be necessary to produce a significant decline in interest rates. Moreover, there is considerable evidence that the marginal efficiency of investment schedule is rather inelastic, with the result that a substantial fall in interest rates will be necessary to stimulate investment significantly. Finally, for reasons indicated above, it is doubtful whether the "net claims effect" on expenditures is strong enough to be of much importance.

Thus, it has come to be generally agreed that the "automatic" forces in the economy tending to keep the system operating at full employment are too weak to be relied upon and that it is therefore necessary to place considerable emphasis on deliberate monetary and fiscal policies of one kind or another to maintain full employment.

It may be well to refer again in this context to the so-called "automatic stabilizers" that were discussed at some length in Chapters 6 and 13.[8] The automatic fiscal stabilizers discussed in Chapter 6 prevent changes in GNP from being reflected fully in changes in disposable personal income; in this way they reduce the size of the multiplier so that, for example, an autonomous decline in expenditures will cause a smaller decline in GNP than would occur in their absence. The automatic monetary stabilizer discussed in Chapter 13 comes into operation when, with a constant stock of money, a decline in income (even with unchanged wages and prices) reduces the transactions demand for money, causing interest rates to fall and imparting a stimulus to investment.

---

[8] See Chapter 6, pp. 134–37, and Chapter 13, pp. 258–59.

These automatic fiscal and monetary stabilizers cut down the degree of instability in the economy. However, they are not capable of preventing changes in GNP entirely, since they are brought into operation as a *result* of changes in GNP. Moreover, while they reduce the magnitude of a recession brought on by a decline in autonomous expenditures, they cannot, by themselves, restore the economy to full employment. Price flexibility, if it is present, is also an automatic stabilizer, but it is more powerful, in a sense, than the stabilizers discussed in Chapter 6 and 13, since it is capable not only of checking a decline but of actually restoring the economy to full employment. This can be seen by reference to Figure 15–1. If the economy is operating at employment N_1, with real wage $(w/p)_1$ and unemployment *AB* (in Panel A) and if wages are flexible, money wages will fall as unemployed workers offer to work at reduced wage rates. As explained earlier, this will reduce costs and prices, thereby increasing the real value of the money stock and causing the *LM* curve (in Panel C) to shift upward from its original position $LM(p_1)$. If the "net claims effect" is operative, the decline in wages and prices will also cause the *IS* curve to shift to the right. This process must continue until the real wage has fallen to $(w/p)_f$ and full employment has been restored, because until this has happened wages and prices will continue to fall, stimulating aggregate demand indirectly through shifts of the *LM* curve, and perhaps also the *IS* curve. Unfortunately, while price flexibility if it were operative would be capable of bringing the economy all the way back to full employment, in practice it cannot be relied upon to achieve this purpose for reasons explained above.[9]

[9] A distinction is sometimes made between *proportional stabilizers,* which generate additional demand equal to some portion of the amount by which income deviates from the full employment level, and *integral stabilizers,* which generate additional demand in proportion to the *cumulative* amount by which income deviates from the full employment level. Proportional stabilizers can reduce the severity of a recession but cannot, by themselves, restore the economy to full employment. Integral stabilizers, on the other hand, are capable of restoring full employment. In this context, the automatic stabilizers discussed in Chapters 6 and 13 are proportional stabilizers, while flexible wages and prices are an integral stabilizer. For a further discussion, see A. W. Phillips, "Stabilisation Policy in a Closed Economy," *Economic Journal,* Vol. LXVI, June 1954, pp. 290–323.

Chapter 16 INFLATION

During much of the period since World War II, inflation has been a major economic problem in the United States. It has been a source of much controversy, the arena of which has ranged all the way from the learned journals in which professional economists publish their studies to the platforms of political parties and the speeches of their candidates. The controversy has had many aspects: What are the causes of inflation? What are its effects on the functioning of the economic system and on the lives of the people? What are the best policies to employ to deal with inflation? To what extent should economic policy be directed at preventing inflation when such prevention interferes with the achievement of other economic and social objectives?

Although we have learned a good deal about inflation in the last 15 years or so, it is only fair to point out that the subject is still fraught with a great many difficulties, and there is room for a considerable amount of legitimate difference of opinion. The student should bear this in mind as he reads the present discussion which attempts to present various viewpoints while at the same time making clear the author's views on the various issues involved.[1]

THE NATURE AND SIGNIFICANCE OF INFLATION

Before turning our attention to the economic analysis of the causes of inflation, let us consider briefly what inflation is and what its economic effects are.

What is Inflation?

Inflation is perhaps best defined as *a tendency toward a continuing rise in the general level of prices.* The reason for including the phrase "tendency toward" in the definition is that it is common to distinguish between *open*

[1] For an intensive and valuable review of the literature dealing with various aspects of inflation, the reader is referred to Martin Bronfenbrenner and F. D. Holzman, "Survey of Inflation Theory," *American Economic Review,* Vol. LIII, September 1963, pp. 593–661.

inflation and *suppressed inflation.* Open inflation, to which we shall confine our attention for the most part, is a situation in which inflationary tendencies are permitted to express themselves in the form of an actual rise in the price level. Suppressed inflation, on the other hand, is a situation in which upward pressures on the price level are present but are prevented from producing an actual rise in prices through the imposition of direct controls in the form of governmentally enforced ceilings on the prices of individual products. Direct price controls are commonly used in wartime as a means of suppressing inflation. Since they tend to interfere with the functioning of the price system as a means of controlling the allocation of resources and of rationing supplies of goods and services among users, they commonly have to be supplemented by systems of priorities, production controls, and rationing of commodities. The use of direct price controls is virtually unavoidable during wartime emergencies, but when used for extended periods in time of peace, such controls are likely to produce distortion and misallocation of resources and reduce economic efficiency so seriously as to make them more undesirable than the inflation they are supposed to prevent.

The phrase "continuing rise" in our definition of inflation is meant to emphasize the fact that inflation involves *rising* prices rather than merely *high* prices. (There is no very meaningful standard—other than an arbitrary historical norm—on which to base the statement that prices are "high.") That is, inflation is a *dynamic process rather than a static state,* and the degree of inflation is capable of being expressed quantitatively as the rate of increase of some selected price index during a specified period of time.

The Economic Effects of Inflation

If all prices in the economy, including the valuations of all assets and all debts, were to move together in the same proportion, changes in the general level of prices would be of no significance. Every individual's money income, the prices of the things he was interested in buying, the value of all of his assets, and the amount he owed under debt contracts would all rise in the same proportion, thus leaving him in the same financial position as before. Such conformity in the behavior of prices is, however, far from being realized in practice. During periods of inflation when most prices are rising, the rates of increases in individual prices vary greatly, and some prices may remain unchanged or even decline.

As a result of these differential movements in the prices of different goods, services, assets and debts, inflation may produce substantial changes in the distribution of income and wealth among families and social classes. The traditional generalizations about the distributional effects of inflation are somewhat as follows: (1) Profit recipients gain from inflation because costs tend to lag behind selling prices in adjustment. (2) Wage earners may gain or lose depending upon the speed with which their wages adjust to changes in prices. (3) Unionization may accelerate the rise in wages and protect workers against the ravages of inflation; on the other hand, in some cases the contractual nature of union wage arrangements may slow down the process of wage adjustment and worsen the position of workers in time of inflation. (4) Salaried workers, such as teachers and white collar groups, are likely to be hurt by inflation, since their salaries are slow to adjust to changes in prices. (5) Recipients of interest and rent are likely to be hurt seriously by inflation, since these payments are often fixed by contracts which are subject to renegotiation only after the lapse of considerable periods of time. (6) Debtors, whose contracts are fixed in terms of money, are benefited by inflation since the real burden of their debt is reduced. (7) Persons holding wealth in the form of common stocks, real estate, and stocks of goods may benefit if the prices of these elements of wealth rise faster than the general level of prices. (8) Persons who hold wealth in the form of currency, bank deposits, or debt claims whose nominal value is fixed in money are hurt by inflation.

Obviously, any particular person may fall in several of these groups, and the net effect of inflation on his position will depend upon the sources from which he receives his income, the form in which he holds his assets, and the extent to which he is in debt. In general, there is some reason to expect that inflation will tend to increase the degree of inequality in the distribution of income and wealth, because profits, which are an especially important source of income for wealthy persons, are supposed to be increased by inflation and because such persons are likely to be quite skillful in investing their accumulated wealth in forms such as common stocks and real estate that respond favorably to inflation.

Several efforts have been made in the last few years to measure the effects of inflation on the distribution of income and wealth.[2] For the most part, these studies have not been particularly successful, chiefly because distribution is affected by many factors, and it has proved to be difficult to disentangle the effects of inflation from the other forces at work. In general, however, it does not seem that mild inflation of the kind we have experienced in the last few years has very strong effects on the distribution of income. Some groups, however, have clearly been affected adversely, the most striking examples being older retired persons dependent on interest

---

[2] G. L. Bach and Albert Ando, "The Redistributional Effects of Inflation," *Review of Economics and Statistics,* Vol. XXXIX, February 1957, pp. 1–13; S. E. Harris, *The Incidence of Inflation: Or Who Gets Hurt?* Study Paper No. 7, Study of Employment, Growth, and Price Levels, Joint Economic Committee, Congress of the United States (1959); and Oswald Brownlee and Alfred Conrad, "Effects upon the Distribution of Income of a Tight Money Policy," in *Stabilization Policies* (Research Studies of the Commission on Money and Credit, 1963), pp. 499–558. For a good brief summary, see *Staff Report on Employment, Growth, and Price Levels,* Joint Economic Committee, Congress of the United States (1959), pp. 110–14.

income, life insurance, annuities, pensions, and the like. Government employees and persons employed by educational institutions and religious and charitable organizations have also been adversely affected. In addition, the postwar inflation seems to have produced a significant increase in the concentration of accumulated wealth, particularly in the hands of the very wealthiest families.

In addition to its effects on the distribution of income and wealth, inflation can affect a nation's balance of payments adversely. This has come to be an important consideration in the United States in the last few years, because we have been troubled with a serious and continuing balance-of-payments deficit. If prices rise more rapidly in the United States than abroad, our products become more expensive compared to those of our foreign competitors, both in the U.S. market and in foreign markets. As a result our imports may increase and our exports decline, thereby aggravating the balance-of-payments problem.

Inflation is often said to have effects on the growth and prosperity of the economy generally, although there is considerable uncertainty concerning the nature of these effects. Two sharply divergent points of view may be recognized. According to the first of these, sound prosperity and steady economic expansion cannot be achieved unless inflation is prevented. Proponents of this view contend that expectations of continuing inflation weaken the propensity to save, thus reducing the resources available for capital accumulation, which is one of the major sources of long-run economic growth and rising living standards. In addition, it is contended that continuing inflation introduces an element of uncertainty into economic calculations and reduces economic efficiency by channeling the energies of business men in the direction of protecting their fortunes against the ravages of inflation at the expense of productive economic activity.[3] The contradictory view is that a continuing gentle increase in the price level has a buoyant effect on economic activity and is conducive to prosperity and healthy economic growth.[4]

While a fairly persuasive case can be made for either of these two sharply contradictory attitudes toward inflation, there is very little solid evidence to support either view. The truth is that it is quite possible to find in different countries and at different times examples of prosperity and rapid economic growth accompanied by rising prices, falling prices, or stable prices.[5] Gen-

---

[3] For an elaborate exposition of the view that a stable price level is a necessary prerequisite for sound prosperity and expansion, see "Creeping Inflation," *Federal Reserve Bank of New York Monthly Review*, June 1959, pp. 86–94.

[4] The late Professor Sumner H. Slichter of Harvard University was a strong advocate of this view.

[5] See *Staff Report on Employment, Growth, and Price Levels, op. cit.*, Tables 1–2 and 1–3, pp. 12–13.

eral economic health is clearly not rigidly tied to any particular kind of price level behavior.

What price index should be used in judging whether inflation is present and in measuring its extent? For reasons discussed in an earlier chapter, it appears that the Consumer Price Index has an upward bias which causes it to exaggerate increases in the price level.[6] Consequently, an all-out effort to stabilize this index would require deflationary policies, which might, for reasons to be explained later in this chapter, be very costly in terms of reduced output and increased unemployment. Moreover, since the Consumer Price Index contains the prices of many services which are not traded internationally, the Wholesale Price Index is a better guage of our international competitive position. On the other hand, the Consumer Price Index has an important effect on movements in money wages, partly because there are often escalation provisions in labor contracts under which wages are automatically adjusted upward when the index rises. The implicit price deflator for the GNP is the most comprehensive index of the prices of newly produced goods and services, but like the Consumer Price Index, it probably contains a significant upward bias. All things considered, there is no single index that can be used as a measure of inflation; indeed, a full understanding of price developments often requires careful study of movements of subdivisions of the Consumer Price Index and the Wholesale Price Index, such as the indexes for food prices, medical services, and industrial commodities.

Types of Inflation

We shall discuss two types of inflation: *demand-pull inflation* and *market-power inflation.* Demand-pull inflation is the classic variety that occurs when the aggregate demand for goods and services exceeds the productive capacity of the economy. In recent years, it has come to be recognized that in a modern industrial economy containing large labor and business organizations, inflation can occur even in the absence of excessive aggregate demand, as a result of the exercise of monopoly power by these organizations. Inflation of this kind we shall refer to as market-power inflation.

DEMAND-PULL INFLATION

Demand-pull inflation has frequently been characterized as a situation in which there is "too much money chasing too few goods." The great inflations which have so often occurred during or immediately following wars have invariably been of this type, and in fact until fairly recently economists have commonly taken the position that all inflationary episodes were basically due to demand pull.

[6] For a discussion of the construction and properties of this and other indexes, see Chapter 4.

A Simple Model of Demand-Pull Inflation

A simple model which will serve to illustrate the essence of demand-pull inflation is presented in Table 16–1.[7] For simplicity, let us assume that the maximum output that can be produced at full employment is $1,000 billion when valued at the prices prevailing prior to the beginning of the inflationary episode. In period one, GNP is $1,000 billion, of which $250 billion is taken in taxes, leaving disposable income of $750 billion. Consumption expenditures amount to 80 percent of disposable income, or $600 billion, investment expenditures amount to 24 percent of disposable income, or $180 billion, and government purchases of goods and services come to $220 billion. Thus, total expenditures (consumption plus investment plus government purchases) amount to exactly $1,000 billion, which is just the value of full employment output at current prices. The economy is, therefore, in full employment equilibrium in period one.

TABLE 16–1

Hypothetical Illustration of Demand-Pull Inflation

Symbol	Meaning	How Calculated	Time Period 1	2	3
y	Full-employment GNP at prices of period 1	Given constant	1,000	1,000	1,000.00
Y	GNP at current prices	$C_t + I_t + G_t$	1,000	1,020	1,040.40
T	Taxes	$0.25Y_t$	250	255	260.10
Y^d	Disposable income	$Y_t - T_t$	750	765	780.30
C	Consumption	$0.8Y_{t-1}^d$	600	600	612.00
I	Investment	$0.24Y_{t-1}^d$	180	180	183.60
G	Government purchases	$0.24Y_{t-1}$ in period 2 and after	220	240	244.80
S_p	Planned (ex ante) saving	$Y_{t-1}^d - C_t$	150	150	153.00
S_a	Actual (ex post) saving	$Y_t^d - C_t$	150	165	168.30
N	Inflationary gap	$Y_t - Y_{t-1}$	...	20	20.40
p	Price index (period 1 = 100)	Y_t/y	100	102	104.04

Suppose that in period two the government increases its purchases of goods and services by $20 billion, to $240 billion. We shall assume that both consumption and investment adjust to disposable income with a lag of one period. As a consequence, both consumption and investment are the same in period two as in period one—that is, consumption is $600 billion and investment $180 billion. However, total expenditures (consumption

[7] For a more sophisticated presentation of a model somewhat similar to the one presented here, see Arthur Smithies, "The Behavior of Money National Income under Inflationary Conditions," *Quarterly Journal of Economics*, Vol. LVII, November 1942, pp. 113–27, reprinted in Arthur Smithies and J. K. Butters (eds.), *Readings in Fiscal Policy* (Homewood, Ill.: Richard D. Irwin, Inc., 1955), pp. 122–36.

plus investment plus government purchases) are equal to $1,020 billion in period two. Since total output at the prices of period one is still only $1,000 billion, the expenditure of $1,020 billion on only $1,000 billion of goods and services must drive prices up in proportion. Thus the price level in period two must rise to 102 percent ($1,020 billion divided by $1,000 billion) of the level in period one, with GNP at the new prices amounting to $1,020 billion.

Let us assume that the tax system is proportional and of such a nature that tax collections are always 25 percent of GNP. Then taxes will rise to $255 billion (25 percent of $1,020 billion) in period two, and disposable income will be $765 billion ($1,020 billion minus $255 billion). Consumption and investment expenditures in period three adjust to the disposable income of period two. Since both prices and money incomes have risen by 2 percent between period one and period two, leaving real disposable income unchanged, we may assume as a first approximation that families consume the same fraction of disposable income and businesses invest the same proportion of disposable income as before. That is, consumption expenditures rise to $612 billion (80 percent of $765 billion), and investment expenditures rise to $183.6 billion (24 percent of $765 billion) in period three. In period two, when it raised its purchases to $240 billion, the government attempted to buy 24 percent of the GNP, which amounted to $1,000 billion at the time the decision to increase expenditures was made. Suppose that the government continues its effort to obtain 24 percent of national output in period three, and in order to do so raises its expenditures to $244.8 billion (24 percent of the previous period's GNP of $1,020 billion). Then in period three total expenditures (consumption plus investment plus government purchases) will amount to $1,040.4 billion. Total output valued at period one's prices is still only $1,000 billion; the expenditure of $1,040.4 billion on $1,000 billion of output must drive the price index up to 104.04 ($1,040.4 billion divided by $1,000 billion). Thus, the price index rises by a further 2 percent in period three (note that $1.0404 \div 1.02 = 1.02$, i.e., 104.04 is 2 percent greater than 102).

As long as the relationships indicated in Table 16–1 continue to hold, the price level will continue to rise at a rate of 2 percent per period. The model of Table 16–1 consists of the following six equations.

$$C_t = c(Y_{t-1} - T_{t-1}) \qquad (1)$$
$$T_{t-1} = vY_{t-1} \qquad (2)$$
$$G_t = gY_{t-1} \qquad (3)$$
$$I_t = k(Y_{t-1} - T_{t-1}) \qquad (4)$$
$$Y_t = C_t + I_t + G_t \qquad (5)$$
$$p_t y = Y_t \qquad (6)$$

The model employs an expenditure lag, which is reflected in the fact that the current period's consumption, investment, and government purchases depend upon the previous period's income. The marginal propensities to consume and to invest out of aftertax income are represented by c and k, respectively; the marginal propensity to pay taxes is represented by v; and the government's propensity to spend on goods and services is represented by g.

The model can readily be solved to show the manner in which the price level in one period depends upon the price level in the previous period. Substituting Equations 1, 2, 3, and 4 into Equation 5, we obtain

$$Y_t = c(Y_{t-1} - vY_{t-1}) + k(Y_{t-1} - vY_{t-1}) + gY_{t-1}$$

or

$$Y_t = [(c + k)(1 - v) + g]Y_{t-1} \tag{7}$$

Substituting Equation 6 into Equation 7, we obtain

$$p_t y = [(c + k)(1 - v) + g]p_{t-1}y$$

and eliminating y from both sides, we have

$$p_t = [(c + k)(1 - v) + g]p_{t-1}$$

For the example worked out in Table 16–1, c is 0.8, k is 0.24, v is 0.25, and g is 0.24. Thus, we have

$$p_t = [(0.8 + 0.24)(1 - 0.25) + 0.24]p_{t-1}$$

or

$$p_t = 1.02p_{t-1}$$

That is, the price level rises at a rate of 2 percent per period.

The essence of this model is very simple. Consumers are trying to buy 60 percent of national output, business firms are trying to buy 18 percent for investment purposes, and government is attempting to obtain 24 percent. Thus the inflation is due to the fact that the aggregate of all demands on national output add up to 102 percent of the output available. Thus, prices are driven up by 2 percent each period. Since the proceeds from the sale of product also constitute income, aggregate gross income also rises by 2 percent each period. The rise in incomes in turn pushes up demand again so that the inflationary process continues. The reader should note that this model is operative only when the economy is at full employment and when aggregate demand adds up to more than 100 percent of available output (i.e., when $(c + k)(1 - v) + g > 1$). When the economy is operating short of full employment, any excess demand operates at least partly to raise output rather than prices, and when aggregate demand falls short of total output (even when the economy is currently at full employment), the shortfall of demand operates at least partly to reduce output rather than prices.

Table 16–1 illustrates the concept of an *inflationary gap.* The inflationary gap in any period is the amount by which aggregate demand in that period exceeds the total available output of goods and services valued at the prices of the preceding period. For example, in period two, aggregate demand is $1,020 billion, which is $20 billion greater than total output available valued at the prices of period one, so that the inflationary gap in period two is $20 billion. By a similar calculation, the gap in period three is $20.4 billion. The gap is a measure of the inflationary pressure; in this example, the gap is always 2 percent of total output valued at the previous period's prices.

The inflationary gap also represents the amount by which the current period's investment plus government purchases exceeds the previous period's saving plus taxes. For example, in period two investment is $180 billion and government purchases amount to $240 billion, for a total of $420 billion, while in period one taxes were $250 billion and planned saving was $150 billion, for a total of $400 billion. Thus, the excess of period two's investment plus government purchases over period one's saving plus taxes amounts to $20 billion, the amount of the inflationary gap for period two. The gap does not, however, show up in the national income statistics. In the model of Table 16–1, it is eliminated partly by a rise in taxes resulting from the rise in income, and partly by the appearance of unplanned saving. For example, in period two, $5 billion of the gap is eliminated by a rise of $5 billion in taxes (from $250 billion in period one to $255 billion in period two). Planned or ex ante saving, which is computed by taking the difference between the previous period's disposable income and the current period's consumption, amounts to $150 billion ($750 billion minus $600 billion). However, actual or ex post saving, as shown in the national income accounts, which is computed by taking the difference between the current period's disposable income and the current period's consumption, amounts to $165 billion ($765 billion minus $600 billion). Thus, the remainder of the inflationary gap in period two is filled by the appearance of $15 billion of unplanned saving. Of course, the gap is reopened by the rise in expenditures in period three—in fact, the gap in that period is larger (by 2 percent), amounting to $20.4 billion.

Some Further Complications

There are a number of factors which affect the process of demand-pull inflation which are not taken into account in the simple model presented in Table 16–1. Some of these factors affect the speed of the inflation, others tend to strengthen the inflationary process, while still others tend to weaken it. We now turn to a brief consideration of some of these factors.

1. In reality inflation is a continuous process of adjustment of prices, incomes, and expenditures, rather than a series of discrete steps such as those

shown in Table 16–1. There is, however, some kind of average lag in the adjustment of expenditures to rising incomes, and the length of this lag is probably the chief factor determining the speed with which prices rise. Unfortunately there is not much that can be said concerning the probable length of this lag, which undoubtedly depends upon many factors that vary from one inflationary situation to another. One factor that may be noted is the expectations of the people concerning the continuation of the inflation. In periods of hyperinflation, such as have sometimes occurred in wartime or in the immediate aftermath of wars, the pace of the inflation may become so rapid that the people expect a significant deterioration in the purchasing power of money from one day to the next. Under these conditions, persons receiving payments of money may hurry to spend them as quickly as possible in order to avoid the loss of purchasing power. Thus, expenditures are adjusted very rapidly to incomes and prices, and the speed of the inflation is greatly accelerated.

2. The extent and rapidity of the adjustment of wages to changes in product prices may affect both the strength and speed of the inflationary process. If the adjustment of wages to prices is incomplete, profits will rise relative to wages. Since the marginal propensity to consume out of wage income is probably greater than the marginal propensity to consume out of profits, this will tend to hold down consumption, with a retarding effect on the inflation. On the other hand, the marginal propensity to spend on investment is undoubtedly greater for profit incomes than for wage incomes, so that the shift from wages to profits will stimulate investment. Thus, it is not clear what the net effect of an incomplete adjustment of wages will be on the inflationary process. If the extent of the wage adjustment is taken as given, however, the speed of the inflation will probably be greater the shorter is the lag in the adjustment of wages to prices.

3. Inflation is often associated in the public mind with an unduly rapid expansion of the money supply. If monetary expansion lowers interest rates and stimulates investment to such an extent that investment demand combined with the other components of demand adds up to more than 100 percent of full employment output, this monetary expansion may indeed be regarded as the proximate cause of inflation. However, the initial impetus to inflation may also come from other sources—for example, an increase in government expenditures (as in the above example) or an autonomous upward shift in consumer demand or business investment demand.

4. Inflation increases the transactions demand for money, since it takes larger cash balances to meet payrolls at rising wages and grocery bills at rising prices. If the stock of money is constant after the initial inflationary impetus (whether this is a result of monetary expansion, an increase in government spending, or some other cause), the increase in the demand for money will raise interest rates, and the rise in interest rates will tend to reduce aggregate demand for goods and services, thus tending to check the

inflation. The strength of this monetary effect will be greater the greater is the income elasticity of the transactions demand for money, since the greater is this elasticity the more money will be drawn into the transactions sphere per unit increase in wages and prices. The strength of the effect will also be greater the lower is the interest elasticity of the demand for money and the greater is the interest elasticity of investment (marginal efficiency of investment). If the interest elasticity of demand for money is low, a sharp rise in interest rates will be necessary to accomplish the needed economization of cash balances, while if the interest elasticity of investment is high, each unit rise in the interest rate will have a rather strong restrictive effect on investment. To summarize, the damping effect of monetary restriction will be stronger (*a*) the greater the income elasticity of the transactions demand for money, (*b*) the smaller the interest elasticity of demand for money, and (*c*) the greater the interest elasticity of marginal efficiency of investment. If the money supply is held constant, the tightening of money automatically produced by inflation will in due course bring the inflation to a halt. But if the elasticity conditions are such as to make the monetary effect weak, a considerable rise in the price level may be possible before the monetary effect stops the inflation. Of course, the monetary effect may be greatly strengthened if the monetary authorities, instead of merely holding the money supply constant, take action to reduce it.

5. Inflation reduces the real value of the stock of net claims against the government, and this may bring about a reduction in private expenditures at given levels of real income, thus helping to check the inflation.

6. If the tax system is progressive rather than merely proportional, as was assumed in the model presented in Table 16–1, an increasing fraction of income is drained away through tax collections as prices rise, and this rise in tax collections can exert a powerful force to check the inflation. In fact, in our economy the progressive federal personal income tax is probably the strongest force working to provide an automatic check to inflation.

7. If there are elements of income (such as salaries, interest, and the like) which remain constant in terms of money (or rise less than in proportion to prices), this may help to stop the inflation by bringing about a redistribution of income, provided the marginal propensity to spend out of these elements of real income is greater than the marginal propensity to spend on the part of the recipients of other elements of income, which must be rising more rapidly than the price level and therefore increasing in real terms. That is, the inflation may be checked by squeezing enough resources out of fixed income classes to meet the demands of other groups in the economy.

8. Instead of remaining constant, as assumed in Table 16–1, real GNP will ordinarily be increasing during a period of inflation, due to the growth of the labor force, the accumulation of capital, and the advance of technology. Thus, the amount of goods and services available to satisfy demands will be increasing. One of the more popular economic clichés is expressed in

the words, "The way to deal with inflation is to increase production." The fact is that increasing production does help to some extent to stem the tide of inflation, but the extent of the help is easily exaggerated. The difficulty is that the increase in real income resulting from the rise in production normally will cause an increase in the demand for consumer goods. This increase in demand is likely to absorb something like three fourths of the increased production, so that only about one fourth of the increased output is available to meet the preexisting excess demand. For example, suppose that, as in period two of Table 16–1, $1,000 billion of goods and services valued at current prices are available to meet monetary demands totaling $1,020 billion, so that the price index will be driven up to 102 ($1,020 billion divided by $1,000 billion). In this circumstance, suppose that output at current prices were to rise by 1 percent, to $1,010 billion, and that the marginal propensity to spend out of additional real income were 75 percent. Then aggregate demand would increase by about $7.5 billion (75 percent of $10 billion) to about $1,027.5 billion, and the price index would rise to about 101.7 ($1,027.5 billion divided by $1,010 billion) instead of 102. If we assume that the time period in Table 16–1 is a quarter of a year, a growth of real output of 1 percent per period is equivalent to 4 percent per year. It is apparent that growth at this rate, which is somewhat greater than the growth we have generally experienced in normal years of full employment, is likely to have only a rather mild antiinflationary effect. It seems clear that increased production is no panacea for dealing with inflation.

9. Inflation will affect the nation's foreign trade position. Rising domestic wages and prices will tend to raise the prices of export products and of domestic goods for which imports can be substituted. If exchange rates are fixed by government intervention in foreign exchange markets, as is the case under the present international monetary system, exports will decline and imports will rise, thus pushing the nation's balance of payments in the direction of a deficit.[8] While such a development may cause a drain on the nation's gold and foreign exchange reserves and create international financial problems, it will help to check the domestic inflation, since fewer goods will be shipped out of the country as exports and more will be brought in as imports, thus increasing the total supply of goods available to meet the excessive demands.

10. Expectations may either accelerate the pace of inflation or help to check it, depending upon their nature. If consumers and businessmen expect the rise in prices, once started, to continue for some time, they may accelerate their purchases of consumer goods, inventories, and fixed capital items in order to obtain desired or needed goods before prices rise further. This kind of reaction will, of course, tend to accelerate the rise in prices. On the other hand, if the prevalent expectations are that the rise in prices will quickly reverse itself—perhaps because there is confidence that the government will

[8] For an extensive discussion of these matters, see Part V.

take prompt and effective measures to check the inflation and that it will be followed by deflation—purchases may be postponed in the expectation of lower prices in the future. Such behavior will itself tend to stop or at least slow down the rise in prices. Which of these two types of expectations will prevail will depend upon circumstances, so it is very difficult to generalize concerning this matter.

Policies for Checking Demand-Pull Inflation

In principle at least, demand-pull inflation can be checked by the use of general monetary and fiscal policies. The above discussion refers to the damping effects produced by the automatic tightening of credit if the supply of money is held constant and by the rise in tax revenues resulting from the rise in money incomes with a progressive tax system. However, monetary and fiscal policies can be used positively to check inflation by taking action to reduce the money supply, to increase tax rates, to reduce government expenditures, or various combinations of these policies. There are, of course, numerous practical problems involved in the effective application of monetary and fiscal policies to deal with demand-pull inflation, and there are important questions related to the choice of the proper combination of policies to be adopted.

MARKET-POWER INFLATION

In a perfectly competitive economy, all prices would be market-determined; that is, no individual buyer or seller would, by himself, be able to change the price of a product or service. In such an economic environment, a general rise in prices could only come about because of excessive aggregate demand, and any inflation that occurred would therefore necessarily be of the demand-pull variety. In a modern industrial economy such as the United States, however, there are many important departures from the conditions of perfect competition, which result from the existence of large business firms and labor organizations that possess substantial monopoly power, enabling them to change prices and wages, at least within limits, by their individual action. In these circumstances, it is possible for inflation to result from the interacting exercise of market power by important groups in the economy even in the absence of excessive aggregate demand. There are a number of behavior patterns by labor and business which, singly or in combination, may produce market-power inflation.

Wages and Labor Productivity

One of the most pervasive characteristics of the American economy throughout nearly all of our history has been the steady rise in money wage rates. Indeed, it is primarily through rising money wages that the fruits of our phenomenal economic progress have been so widely shared among all classes in our society. But, while rising wages may be a constructive element in the process of economic growth, wage rates can, under some circumstances, rise so rapidly as to imperil the stability of prices.

How rapidly money wages can rise without putting upward pressure on prices depends primarily on what is happening to the productivity of labor. This can perhaps best be illustrated by means of a simple example. Suppose in a particular year, using the machines and techniques then available, the workers employed by a particular firm are able to turn out 10 units of that firm's product per hour. If these workers are paid a money wage of $2 per hour, the labor cost per unit of output for that firm is $0.20 ($2 divided by 10 units). Now suppose that in the following year, as a result of the installation of more and better capital equipment, more skilled management, and possibly better training or more effort on the part of the workers, hourly output per worker rises by 10 percent, to 11 units per hour. Under these conditions, if the money wage rate also rises by 10 percent to $2.20 per hour, labor cost per unit of output will continue to be $0.20 per unit ($2.20 divided by 11 units). Thus, while *wage rates* have risen, *labor costs* per unit of the product remain the same as they were in the first year.

Thus, a rise in money wages that merely parallels a rise in the productivity of labor does not raise the labor cost of producing output. On the other hand, if money wages rise more rapidly than productivity, labor costs are pushed up. Suppose, in the above example, instead of merely rising by 10 percent between the first and the second year, the money wage rate had risen from $2 to $2.40, an increase of 20 percent. In this case, the labor cost per unit of output would have been $0.218 ($2.40 divided by 11 units) in the second year, a rise of 9 percent over the $0.20 unit cost that prevailed in the first year.[9]

The above analysis can be extended, at least in a crude way, to the economy as a whole. If the economywide average increase in money wage rates exceeds the average increase in output per worker, there will be a general increase in labor costs per unit of output. And, as will be explained more fully below, such a rise in labor costs will be likely to exert upward pressure on the general price level of goods and services.

The Relation between Money Wages and Unemployment

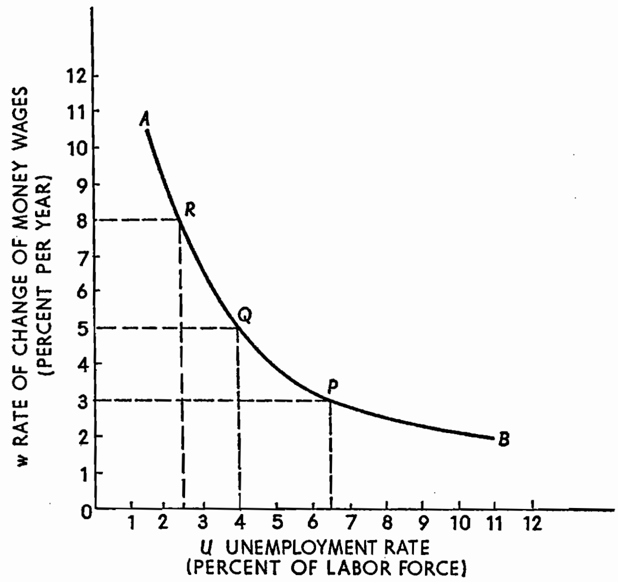
Since the rate of increase in money wages relative to the increase in labor productivity has an important effect on unit labor costs and therefore on

[9] It should be noted that so-called "fringe benefits" should be included, along with actual money wage payments, in the calculation of unit labor costs. Fringe benefits include such things as vacation pay, payments by the employer into pension and retirement funds, employer contributions to social security programs, and so on. Fringe benefits have become an increasingly important element in labor costs in recent years.

prices, the forces affecting money wages are of crucial importance in the analysis of inflation. There is considerable evidence that the rate of increase in money wages for the economy as a whole increases as the rate of unemployment declines. This relationship, of the type depicted by curve *AB* in Figure 16–1, has come to be known as a "Phillips curve," because it was first discovered in a study of the British economy by A. W. Phillips.[10] In the case depicted in Figure 16–1, wages rise at a rate of 3 percent per year when the unemployment rate is 6.5 percent (point *P*). If unemployment falls to 4

FIGURE 16–1

Phillips Curve



percent, the rate of increase in wages is 5 percent per year (point *Q*); if unemployment declines further to 2.5 percent, wages rise at 8 percent per year (point *R*); and so on. Although the reasons for this inverse relation between the unemployment rate and the rate of increase in money wages have not been conclusively established, several explanations are possible.[11]

The Role of Collective Bargaining. It seems very likely that one of the factors underlying the Phillips curve relationship is the collective bar-

[10] A. W. Phillips, "The Relation between Unemployment and the Rate of Change of Money Wage Rates in the United Kingdom, 1861–1957," *Economica,* New Series, Vol. XXV, 1958, pp. 283–99.

[11] For a valuable discussion of the Phillips curve as applied to the U.S. economy, see P. A. Samuelson and R. M. Solow, "Analytical Aspects of Anti-Inflation Policy," *American Economic Review,* Vol. L, May 1960, pp. 177–94, reprinted in W. L. Smith, and R. L. Teigen (eds.), *Readings in Money, National Income, and Stabilization Policy* (2d. ed.; Homewood, Ill.; Richard D. Irwin, Inc., 1970).

gaining process. When unemployment is low and markets for labor throughout the economy are relatively tight, labor unions are quite aggressive in pressing for large increases in money wages. On the other hand, when unions find heavy unemployment among their membership, they are generally somewhat less vigorous in seeking wage increases. Moreover, at times when unemployment is relatively low, markets for goods of various kinds will ordinarily be strong. Under these circumstances, business firms will be reluctant to accept strikes which will require them to curtail production at a time when business is highly profitable. Accordingly, employers will be less vigorous in resisting labor demands for wage increases under such conditions than they will when markets for their products are relatively weak. Thus, the relative bargaining strength of labor and management seems to vary substantially as the rate of unemployment and the general level of business activity changes. Such shifts in the balance of power in collective bargaining negotiations would produce an inverse relation between the rate of change in money wages and the rate of unemployment, as depicted by the Phillips curve.

Structural Imbalances. Another factor that may help explain the existence of a Phillips curve relationship is the occurrence of imbalances between supply and demand in particular labor markets. As the demand for labor expands under the impetus of a rise in aggregate demand, the occupational and skill requirements and the geographic location of the additional labor demanded may not match these characteristics of the unemployed portion of the labor force. As a result, the increasingly frequent appearance of labor shortages may cause the rate of increase in money wages to accelerate as the overall unemployment rate declines. If markets were thoroughly competitive, wages would tend to fall in other occupations and geographic areas in which there was excess labor, and the general level of labor costs would not rise unless excessive aggregate demand was pulling up both product prices and wage rates. Thus, if shortages of labor in *particular* markets cause the *general* level of labor costs to rise in a period of substantial unemployment, it must be due to market power or market imperfections which prevent wages from falling in markets where there is an excess supply of labor.

Generalized Excess Demand for Labor. The factors mentioned above have not satisfied some observers as an explanation of the inverse relation between the unemployment and the rate of change in money wages. In particular, the explanation based on changes in the relative bargaining power of unions and employers does not seem entirely convincing, because only a relatively small fraction of the U.S. labor force is unionized. Of course, wage increases obtained by unions have an effect on wages in nonunionized sectors of the economy, because of the existence of competition for labor and because nonunion employers may often increase the wages of their workers in line with those of unionized workers in order to ward off threats of unionization. However, in periods when wages rise sharply, the wages of nonunion workers often rise earlier and more rapidly (at least in the early stages) than those of union workers. This pattern is obviously hard to reconcile with the bargaining power theory.

An alternative explantion of the inverse relation between the rate of change of money wages and the unemployment rate has been advanced by Richard G. Lipsey.[12] For simplicity, let us assume that there is only one homogeneous type of labor and that all workers receive the same wage. The demand for and supply of labor might be as depicted in Panel A of Figure 16–2. In this case, the equilibrium wage rate is *OA*, since at this wage rate demand and supply are equal. Thus, if the wage rate is *OA* and employment is *OF*, the wage rate will remain unchanged (unless there is a change in demand or supply conditions). Let us, however, consider situations in which the wage rate is below equilibrium. For example, if the wage is *OB*, the supply for labor is *OE* and the demand is *OG*. Thus, there is excess demand of *EG*, and this excess demand will exert upward pressure on the wage, pushing it toward equilibrium.

Now let us suppose the percentage rate at which wages rise varies directly with the excess demand for labor. That is, if the wage is *OC* with excess demand *DH*, the wage will rise faster than will be the case when the wage is *OB* with excess demand *EG*, since *DH* is greater than *EG*. Such a relationship may be expressed by the following equation

$$\frac{\Delta W}{W} = a\frac{D - S}{S} \tag{8}$$

where $\Delta W/W$ is the percentage rate of change of wages. D is the demand for labor, S is the supply of labor, and a is a constant factor of proportionality. That is, the rate of increase in wages is simply proportional to the excess demand for labor expressed as a proportion of the labor supply.

Merely because the labor market is in equilibrium at wage rate *OA*, it does not follow that no one is unemployed. If N is the number of workers employed, V is the number of vacant jobs that employers would like to fill (which we shall simply call "vacancies"), and U is the number of unemployed workers seeking jobs, we have

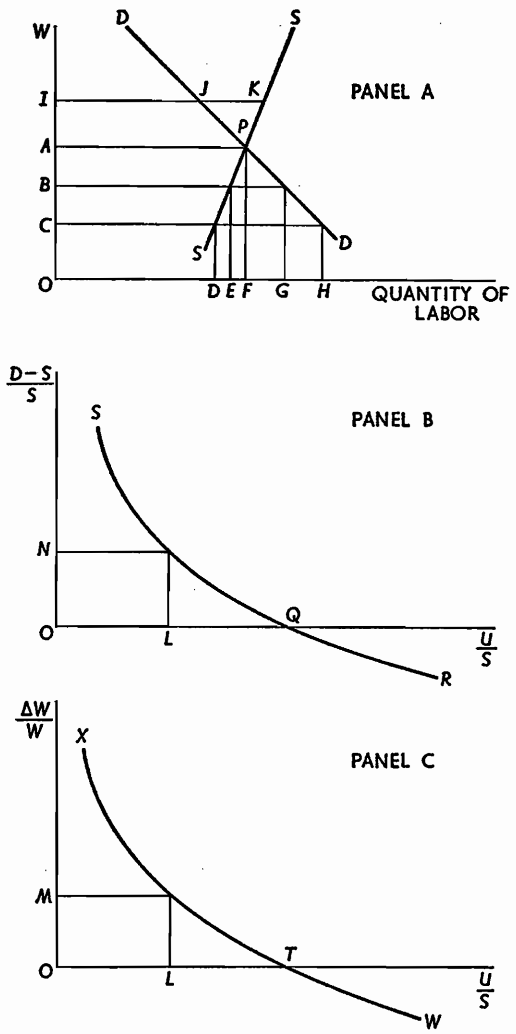
$$D = N + V$$
$$S = N + U$$

That is, the demand for labor is equal to employment plus vacancies, and the supply of labor is equal to employment plus unemployment. Vacancies and unemployment can exist side by side because some workers will want to

[12] R. G. Lipsey, "The Relation between Unemployment and the Rate of Change of Money Wage Rates: A Further Analysis," *Economica*, New Series, Vol. XXVII, February 1960, pp. 1–31, reprinted in R. A. Gordon and L. R. Klein (eds.), *Readings in Business Cycles* (Homewood, Ill.: Richard D. Irwin, Inc., 1965), pp. 456–87.

FIGURE 16–2

Relations between Excess Demand for Labor, Unemployment, and Wage Changes



change jobs and it takes a finite amount of time to find a new job. When the wage is equal to OA in Panel A of Figure 16–2, we have

$$D = S$$
$$V = U$$

That is, when the demand for labor equals the supply, the number of job vacancies will be equal to the number of workers unemployed.

In general, the amount of excess demand can be expressed as

$$\text{Excess demand} = D - S = N + V - (N + U) = V - U$$

When the wage is below the equilibrium level (OA in Panel A of Figure 16–2), excess demand is positive, and vacancies exceed unemployment. Conversely, when the wage is above equilibrium, excess demand is negative—that is, supply exceeds demand—and unemployment exceeds vacancies. This is the situation, for example, when the wage is OI in Panel A of Figure 16–2: in this case, excess demand is negative to the extent of JK.

From the above explanation, it is apparent that the greater the amount of excess demand, the easier it will be to find jobs and the shorter will be the amount of time taken in moving between jobs. Consequently, the greater is excess demand, the smaller will be the amount of unemployment. Panel B of Figure 16–2 shows the kind of relation that would exist between excess demand and the unemployment rate (which is defined as the ratio of unemployment to the supply of labor, U/S). When excess demand equals zero (which corresponds to the equilibrium point, P, in Panel A), the unemployment rate is OQ. As excess demand becomes negative in Panel B (corresponding to values of the wage above equilibrium in Panel A), the unemployment rate rises along the line segment QR. On the other hand, as excess demand rises above zero, the unemployment rate declines along the path QS. Since unemployment cannot become negative, the segment QS is drawn as a curving line which comes closer to the vertical axis as excess demand $[(D - S)/S]$ becomes larger but never reaches the axis. The relation between excess demand and the unemployment rate which is depicted by the curve SQR in Panel B can be expressed as an equation

$$\frac{D - S}{S} = f\left(\frac{U}{S}\right) \tag{9}$$

Substituting from Equation 9 into Equation 8, we obtain

$$\frac{\Delta W}{W} = af\left(\frac{U}{S}\right) \tag{10}$$

The relation in Equation 10 is depicted graphically in Panel C of Figure 16–2. Since the right-hand side of Equation 10 is simply the right-hand side of Equation 9 multiplied by a constant, a, the curves SQR in Panel B and XTW in Panel C will be similar in shape. At point T in Panel C, excess demand will be equal to zero (corresponding to point P in Panel A). As we move to the left from point T, excess demand becomes increasingly greater (according to the relation in Equation 9), and wages rise more rapidly (according to the relation in Equation 8). It is apparent that the curve XTW in Panel C represents a relationship of the Phillips curve type.

A commonsense explanation of the Lipsey theory of the Phillips curve can be given as follows: if those responsible for economic policy want to achieve an unemployment rate less than *OT* in Panel C of Figure 16–2, they must create a positive excess demand for labor, and this will cause wages to rise. For example, if they want to achieve unemployment rate *OL,* they must create excess demand equal to *ON* in Panel B, and this will cause wages to rise at a rate equal to *OM* in Panel C. Moreover, if they want to hold the unemployment rate at that level, they cannot permit the excess demand to be eliminated, but rather must continue to stimulate the demand for labor by fiscal and monetary policies sufficiently to keep the excess demand gap open. If they do this, wages will *continue to increase* at the rate that is associated with that gap.

There is, of course, no reason why one must choose among the three explanations of the inverse relation between the unemployment rate and the rate of change of wages discussed above. All three causes may be operating at the same time—that is, as unemployment declines wages may rise more rapidly, partly because of a shift in the balance of power in collective bargaining, partly because of increasing structural imbalances affecting particular labor markets, and partly because lower unemployment is associated with a generalized increase in excess demand for labor.

The Wage-Price Spiral

A persistent tendency for labor costs to rise because money wages increase more rapidly than the productivity of labor will in due course cause businessmen to pass along the cost increases to their customers through increases in product prices. Such increases may come about through the impersonal working of demand and supply if product markets are quite competitive. In the case of firms selling their products in semimonopolistic markets, the increases may be the result of either (*a*) carefully calculated marginal adjustments of prices to changing costs for the conscious purpose of maximizing profits, or (*b*) the use of some such rule of thumb as the application of a more or less stable percentage markup to unit labor and material costs. But, whatever the process by which the adjustments are made, a persistent tendency for wage increases on an economywide basis to exceed increases in the productivity of labor is bound to lead to a continuing rise in the price level even though there is substantial unemployment.

Although some sectors of the economy may not be directly affected because wage increases in these sectors do not exceed productivity increases, even these sectors are likely to become indirectly involved in the wage-price spiral. This is because products produced in sectors in which prices are being forced up, due to wage increases which exceed productivity increases, will, in many cases, enter as inputs into the production process in other sectors. As a result, material costs in these other sectors will rise, and this in due course will tend to force up final product prices. Thus, inflation resulting from excessive wage increases in a few key sectors of the economy may spread rapidly to produce the phenomenon of general price inflation.

Inflation of the type just described may be still further aggravated to the extent that there is a tendency for wages to be adjusted upward to compensate for increases in the cost of living. Such adjustments may come about either through the operation of so-called escalator clauses in union contracts, which may call for an upward adjustment of money wage rates each time the cost of living index rises by some specified number of percentage points;[13] or, even in the absence of such escalator clauses, prior increases in the cost of living may be used as a basis for negotiating larger wage increases at the time regular contract settlements take place.

Empirical Evidence on the Phillips Curve

Much empirical work has been done in recent years on the determinants of the level of and changes in money wages in the United States.[14] Most of the studies show a clear inverse relation between the rate of change of money wages and the rate of unemployment, although there is still some uncertainty about the exact form of the relationship. It will suffice for our purposes and give the student the general flavor of the empirical work to review the results of one of the major studies, that of George L. Perry.[15] After experimenting with a variety of formulations Perry settled on the following equation, fitted by regression techniques to quarterly data for the period from the fourth quarter of 1947 to the third quarter of 1960:[16]

$$w = -4.313 + 14.711 \frac{1}{U} + 0.367p + 0.424P \qquad (11)$$

---

[13] The contracts between the United Automobile Workers and the major automobile manufacturers have contained escalator clauses since 1948. The clauses call for periodic adjustments in wage rates keyed to changes in the Consumer Price Index. Since their adoption in the automobile industry, escalator clauses have been included in many other collective bargaining contracts, and several million workers are covered by them.

[14] There is an extensive literature on the relation between money wages and unemployment, much of which includes empirical work. Especially worthy of mention are Otto Eckstein and Thomas Wilson, "The Determinants of Money Wages in American Industry," *Quarterly Journal of Economics,* Vol. LXX, August 1962, pp. 379–414; W. G. Bowen and R. A. Berry, "Unemployment Conditions and Movements of the Money Wage Level," *Review of Economics and Statistics,* Vol. LXV, May 1963, pp. 163–72; Edwin Kuh, "A Productivity Theory of Wage Levels—an Alternative to the Phillips Curve," *Review of Economic Studies,* Vol. XXXIV, October 1967; and E. S. Phelps, "Money-Wage Dynamics and Labor-Market Equilibrium," *Journal of Political Economy,* Vol. LXXVI, July/August 1968, Part II, pp. 678–711.

[15] G. L. Perry, *Unemployment, Money Wage Rates, and Inflation* (Cambridge: Massachussetts Institute of Technology Press, 1966).

[16] Perry's original equation (*ibid.,* p. 50, Equation 3.8) also includes a term for the quarterly change in profits. Since we are interested only in the equilibrium relationship, however, we have dropped this term.

Here w is the percentage rate of change per year in hourly wages in manufacturing; U, the unemployment rate (percent); p, the percentage rate of change per year in the Consumer Price Index (CPI); and P is the profit rate in manufacturing (percent; profits after taxes divided by stockholders' equity). That is, money wages in manufacturing rise more rapidly (*a*) the lower is the unemployment rate, (*b*) the faster the CPI is rising, and (*c*) the higher is the profit rate.

In order to isolate the relation between wage changes and unemployment, let us suppose that changes in the CPI are determined by changes in money wages and changes in productivity (output per man-hour); that is

$$p = w - q \tag{12}$$

where q is the rate of increase in output per man hour. For example, if output per man-hour rises by 3 percent per year while money wage rates rise by 4 percent per year, labor cost per unit of output rises by 1 percent per year; Equation 12 assumes that prices would be raised in proportion to the increase in unit labor cost, or by 1 percent per year. Substituting Equation 12 into Equation 11 we have

$$w = -4.313 + 14.711\frac{1}{U} + 0.367(w - q) + 0.424P \tag{13}$$

Solving explicitly for w (which appears on both sides of Equation 13), we obtain[17]

$$w = -6.814 + 23.240\frac{1}{U} - 0.580q + 0.670P \tag{14}$$

This equation shows that a high rate of productivity increase helps to hold down the rate of increase in wages, because it reduces the effect of any given rate of wage increase on unit labor cost, thereby lessening the unfavorable "feedback" effect of increases in the CPI on wages.

To get an idea of the implications of Equation 14 for the shape and position of the Phillips curve, let us make the following assumptions: (*a*) productivity increases at a rate of 3 percent per year ($q = 3$), which is close to the long-run trend rate of increase experienced in the United States; (*b*) the aftertax profit rate in manufacturing is 12 percent ($P = 12$), which is a reasonable assumption for a period of prosperity. Substituting $q = 3$ and $P = 12$ in Equation 14 and simplifying, we obtain

$$w = -0.514 + 23.24\frac{1}{U}$$

[17] *Ibid.*, p. 61, Equation 3.14.

Based on this equation, Table 16–2 gives the rates of increase in wages (w) for selected values of the unemployment rate (U). If these values are plotted graphically, they yield a Phillips curve very similar to AB in Figure 16–1. Note that in Table 16–2, 6.61 percent is the unemployment rate at which money wages rise at the same rate (3 percent) as productivity. An increase in the profit rate (P) will shift the Phillips curve upward to the right, while an increase in the rate of productivity growth (q) will shift it downward to the left.

One variable not mentioned earlier appears in Perry's equation: the profit rate. This is reasonable; an increase in the profit rate means that business can "afford" to increase wages, and it also provides an argument that unions can

TABLE 16–2

Relation between the Unemployment Rate and the Rate of Change of Money Wages*

U *Unemployment Rate (Percent)*	w *Rate of Change of Money Wages (Percent)*
2.00	11.1
3.00	7.2
4.00	5.3
5.00	4.1
6.00	3.4
6.61	3.0
7.00	2.8
10.00	1.8

* Calculated from the equation $w = -0.514 + 23.240 \frac{1}{U}$

use in support of their demands. The profit rate bears at least a loose connection with the unemployment rate; that is, when unemployment is low, business will be good and the profit rate is likely to be relatively high. Actually, this relation between the profit rate and the unemployment rate should be taken into account in calculating the probable equilibrium relationship between the unemployment rate and the rate of change of wages. For example, in Table 16–2, a 4 percent unemployment rate is associated with a 5.3 percent per year increase in wages. If the assumed 12 percent profit rate is associated with a 4 percent unemployment rate, a lower profit rate—say, 11 percent—would presumably be associated with a 6 percent unemployment rate. Assuming, as before, that $q = 3$ percent, but changing P from 12 percent to 11 percent, Equation 14 yields a rate of increase in wages of 2.7 percent for an unemployment rate of 6 percent rather than the increase of 3.4 percent shown in Table 16–2. Thus, it can be seen that if an appropriate adjustment could be made for the relation between the unemployment rate and the profit rate, the Phillips curve would become somewhat steeper than it is in the absence of such an adjustment. However, while there is undoubtedly a loose relation between the unemployment rate and the profit rate, it is scarcely dependable enough to permit a systematic adjustment to be made.

The equation $p = w - q$ (Equation 12, which is used to derive Equation 14 from Equation 11) is no more than a very crude approximation. The CPI includes food prices and the prices of many consumer services and is therefore affected by many forces other than the behavior of wages and productivity in manufacturing industries.

The Perry study is useful in giving a general idea in quantitative terms of the relation between the unemployment rate and the rate of change of money wages. The student is cautioned, however, not to take the equation too seriously. The study covers only the manufacturing sector of the economy, and it is necessary to employ some crude approximations, referred to above, to derive a Phillips curve from it. Moreover, while the "fit" of Equation 11 is very good as time-series regression equations go, there is nevertheless a significant amount of variation in the rate of change of money wages that is left unexplained.

Controversial Issues Relating to the Phillips Curve

While the idea of a Phillips curve showing the relation between the unemployment rate and the rate of change of wages is fairly generally accepted by economists, there is some uncertainty about the stability of the relationship over time. Some economists believe the rate of change of wages (and prices) associated with a given rate of unemployment depends on the time path followed by the economy in reaching that rate of unemployment. One view is that a rapid reduction in unemployment caused by a very sharp increase in aggregate demand will generate bottlenecks in labor (and product) markets, because supply adjustments take time to carry out and supply is therefore often very inelastic in the short run. Thus, when unemployment falls rapidly, there may be a greater increase in wages (and prices) than would have occurred if the same decline in unemployment had occurred more gradually.[18] A similar but slightly different idea is that a given average unemployment rate will be associated with a higher average rate of increase of wages (and prices) over a period of several years if the economy moves up and down in a series of "fits and starts"—as was the case, for example, in 1965–69—than if it grows steadily without significant fluctuations.[19]

---

[18] For a development of this view as applied to the period 1965–68, see W. L. Smith, "High Employment Level without Price Inflation," *Commercial and Financial Chronicle,* October 31, 1968, pp. 20–21.

[19] This view is developed in S. H. Hymans, "The Trade-Off Between Unemployment and Inflation: Theory and Measurement," Smith and Teigen, *op. cit.*

Recently, some economists have contended that the conventional Phillips curve of the type depicted in Figure 16–1 is merely a transitory short-run relationship.[20] In their view, both the demand for labor by employers and the supply of labor forthcoming from workers depend basically on the real wage, that is, the money wage deflated for price changes. There is a so-called "natural rate of unemployment" that is consistent with stable prices or with prices rising or falling at a constant rate. This natural rate of unemployment is the rate corresponding to that real wage at which the demand for and supply of labor are equated. It is possible temporarily to reduce unemployment below the natural rate by expanding aggregate demand. By causing prices to rise, this lowers the real wage, inducing employers to expand production and employment. At first workers do not realize that the increase in prices has reduced their real wage. Once they do realize it, however, they demand and obtain higher money wages, and the reduction in real wages below the equilibrium level can only be maintained by a further rise in prices. As prices and wages continue to rise at a steady rate, workers eventually come to anticipate the price increases and to take these anticipations into account in making their wage demands. Thus, in the course of time real wages and employment fall back to their equilibrium levels, unemployment is restored to its natural rate, and all that is happening is that prices and wages are rising together with real wages remaining constant (or rising as fast as productivity is increasing). The only way to keep unemployment below its natural rate over a long period of time is to repeat the above process over and over—that is, to accept an accelerating rise in prices (for example, 2 percent per year at first, then 3 percent, 4 percent, and so on).

According to this view of the inflation process, the long-run Phillips curve is simply a vertical line drawn at the natural rate of unemployment, as depicted in Figure 16–3. The natural rate may vary from country to country and from time to time in the same country, depending upon the degree of labor mobility, the extent of unionization, and other factors. It may be reduced—i.e., shifted to the left in a diagram such as Figure 16–3—by measures which improve the functioning of labor markets.

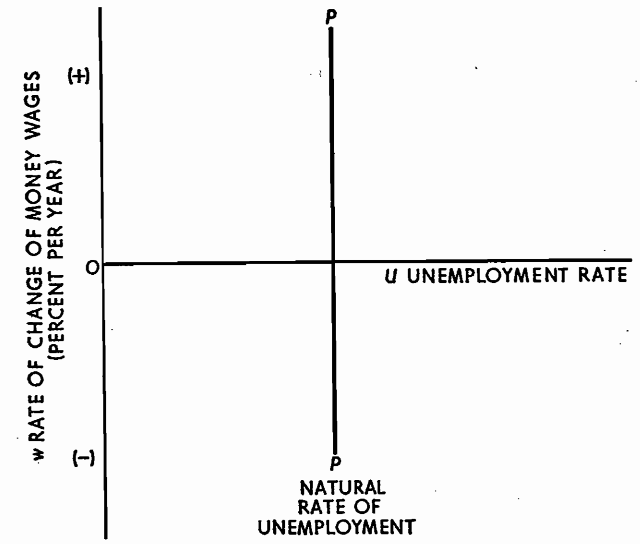
This view of the long-run relationship between inflation and unemployment is based almost entirely on deductive reasoning from certain assumptions about the behavior of employers and workers. Theories that are supposed to hold true only in the long run are exceedingly difficult to subject to any kind of empirical tests, because underlying conditions are unlikely to remain constant long enough to permit the long-run relationships to manifest themselves. And, in the case of this particular hypothesis, there are a number of reasons for believing that the time period required for it to become operative may be very long indeed. There are almost certain to be fluctuations in the strength of private demand, and fiscal and monetary

[20] See Phelps, *op. cit.;* and Milton Friedman, "The Role of Monetary Policy," *American Economic Review,* Vol. LVIII, March 1968, pp. 1–17, reprinted in Smith and Teigen, *op. cit.*

policies are very unlikely to be able to offset the effects of these fluctuations with complete efficiency. Consequently, both the unemployment rate and the rate of change of wages and prices are virtually sure to be subject to variation. Moreover, even if, by some miracle, the Consumer Price Index should rise at a perfectly steady rate for several years, this behavior could only be the result of offsetting changes in the constituent prices that make up the index. Since individual prices are affected by shifts in demand and

FIGURE 16-3

Hypothetical Long-Run Phillips Curve
When Inflation Is Fully Anticipated



supply in particular markets, some prices would rise more rapidly than the average while others were rising less rapidly or even falling. Because of differing tastes, the relative importance of various goods and services in consumer market baskets varies from one consumer to another. Thus, even with the CPI rising at a steady rate, the price levels relevant to particular consumers would be changing in a variety of different, and probably quite erratic, ways. Under such conditions, a policy that was mildly inflationary on the average might continue for a number of years before it became translated into a meaningful consensus of inflationary expectations and thereby lost its power to hold down the rate of unemployment.

Further Aspects of Market-Power Inflation

In the above discussion, the basic cause of inflation is indicated to be a tendency for money wage increases to exceed increases in labor productivity.

This causes unit labor costs to rise, and the rise in costs is passed through into prices. This process is enough to produce a continuous rise in the price level as long as wage increases continue to exceed productivity increases. The inflationary process is further accentuated if the rise in the price level feeds back to cause still further increases in money wage rates through escalator clauses or other aspects of the collective bargaining process.

Inflation of the market-power variety can be caused by factors other than a tendency for money wage rates to rise more rapidly than labor productivity.[21] Indeed, at least three other possible causes can be distinguished.

1. Market-power inflation can be set off, even in a situation in which money wage increases merely match increases in productivity, if businessmen attempt to increase their profits by raising their profit margins, that is, by raising their sales prices relative to unit costs at given levels of output. This may happen in sectors which are characterized by so-called "administered prices." These are industries in which some quasi-monopoly power exists, which gives the businessman some control over the price charged for his product. If the increase in profit margin is a once-for-all increase, it can set off an inflationary spiral, provided wages are adjusted upward to compensate for increases in the cost of living. However, in the absence of a basic tendency for wages to rise more rapidly than the productivity of labor, an inflationary spiral set off by a once-for-all increase in profit margins will die out gradually as wages and prices are brought into a new equilibrium relationship. Only a continuing effort to produce successive increases in the profit margin is likely to lead to an unending inflationary spiral. For this reason, autonomous increases in profit margins do not seem to be a very likely cause of a continuing fundamental problem of inflation.

2. An autonomous increase in raw material prices could also set off an inflationary spiral. Again, however, merely a once-for-all increase in raw material prices would be unlikely to cause an unending upward spiral in the price level. Only a continuing tendency for raw material prices to rise more rapidly than the prices of finished goods would be likely to produce this result. A rise in material prices could come about either through the appearance of bottlenecks in the production of domestically produced raw materials or from a rise in prices in foreign countries which caused imported raw material prices to rise.

3. More generally, a species of market-power inflation could be brought about during a period of economic expansion, even while labor and plant capacity were unemployed in many sectors of the economy, if the expansion of demand was especially rapid in some sectors so that bottlenecks began to appear early in the expansion process. Price increases in sectors where such bottlenecks occurred might be passed through to other sectors where the

[21] Inflation caused by excessive increases in money wages at a time of substantial unemployment is often called "cost-push" inflation or "wage-push" inflation. However, the present writer prefers the more general term, "market-power" inflation.

products of the bottleneck sectors entered as inputs, thus driving up costs and prices in these other sectors. If the sectors in which excess capacity existed were fully competitive, prices would fall in these sectors, thereby keeping the general level of prices from rising significantly. But the existence of market power may enable producers in these sectors to avoid reducing their prices. Another related species of inflation is so-called "demand-shift" inflation. Such inflation could occur even though there was no change at all in aggregate demand and total output and even though the economy was operating substantially short of full employment, provided there was a shift in demand—that is, a decline in demand in some sectors accompanied by an increase in demand in other sectors. Since the existence of market power makes prices more rigid in a downward than in an upward direction, prices probably would tend to decline little in the sectors where demand dropped but might increase fairly sharply in sectors in which demand increased. The price increases in sectors experiencing increased demand might then spread through the whole economy as products produced in those sectors entered as inputs in other industries.[22] However, while general bottleneck or "demand-shift" inflation might cause troublesome temporary problems, it seems unlikely that inflation originating from such causes would continue indefinitely.

Thus, while temporary inflationary tendencies can undoubtedly arise from the three factors referred to above, the basic source of market-power inflation appears to be the tendency for money wage increases to outrun increases in productivity.

Market-Power Inflation and Aggregate Demand

In the above discussion of market-power inflation, little has been said about the behavior of aggregate demand. However, continuing market-power inflation does require that total demand expand at a sufficient pace to absorb the output produced at continually higher prices. There is ample evidence that inflation resulting from a tendency for money wage increases to outrun increases in productivity can be brought to a halt by a sufficient contraction of aggregate demand. Restrictive monetary and fiscal policies will slow the pace of wage increases by reducing excess demand in labor markets and by shifting the balance of collective bargaining against unions and in favor of employers.

As will become apparent from the analysis presented in Chapter 17, in a growing economy it may not be necessary to have an actual decline in aggregate demand in order to restrain inflationary pressures. With a steadily growing labor force and rising productivity, the real GNP that can be

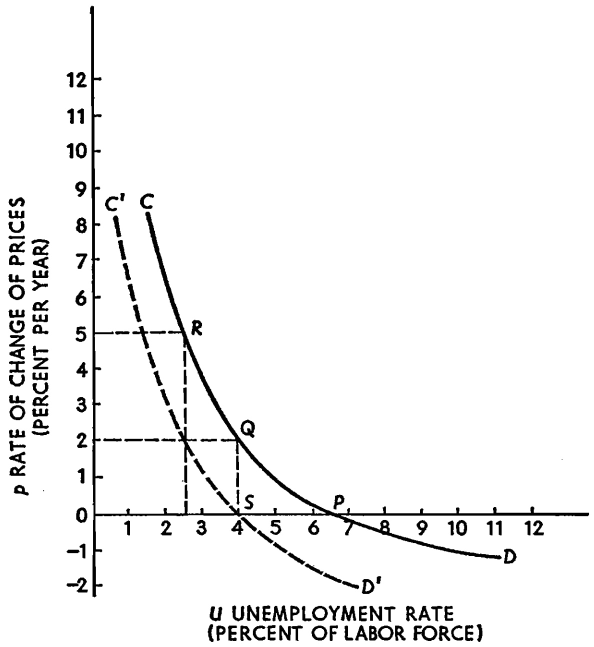
[22] The concept of demand-shift inflation is developed and used to explain price movements in 1955–57 in C. L. Schultze, *Recent Inflation in the United States,* Study Paper No. 1, Study of Employment, Growth, and Price Levels, Joint Economic Committee, 1959.

produced at a given unemployment rate increases from quarter to quarter. Thus, restrictive monetary and fiscal policies which merely cause the growth of real aggregate demand to slow down—without producing an actual decline—may bring about a rise in unemployment, weakening the upward pressures on wages and prices.

It may be noted that even in the absence of any positive action on the part of government to restrict aggregate demand, there are forces that work automatically to produce such restriction. One of these is a progressive tax system which absorbs an increasingly large share of the additions to money incomes as such incomes rise, thereby slowing down the rise in disposable incomes and imposing a check on private spending. A second is the increase in the transactions demand for money which occurs at any given level of income as prices and wages rise. This increase in the transactions demand for money draws a portion of the money supply out of the asset sphere and thereby tends to raise rates of interest. The rise in interest rates, in turn, tends to reduce investment spending and thereby impose a check on aggregate demand. This monetary effect will be more important and powerful in bringing the inflation to a halt (*a*) the greater is the income elasticity of transactions demand for money, (*b*) the smaller is the interest elasticity of the demand for money, and (*c*) the greater is the interest elasticity of investment demand. If they are permitted to operate, these automatic restrictions on aggregate demand must eventually bring market-power inflation to a halt, but a considerable rise in the price level may have to occur before they produce this result. The government can, of course, reinforce these automatic tendencies by adopting positively restrictive fiscal and monetary policies—which would involve such measures as reductions in government expenditures, increases in tax rates, and actions by the Federal Reserve to curtail the supply of money and credit.

While it is possible to put a stop to market-power inflation by a sufficiently vigorous restriction of aggregate demand through fiscal and monetary policy, such an objective can be accomplished only at the cost of a reduction in total output—or at least a slowdown in its growth—and an increase in unemployment. The relationship between inflation and unemployment can be depicted by an adaptation of the Phillips curve. The modified Phillips curve, line *CD* in Figure 16–4, shows the relation between the unemployment rate and the rate of change of *prices,* whereas the original Phillips curve (Figure 16–1) shows the relation between the unemployment rate and the rate of change of *wages.* Since the rate of change of wages can ordinarily be expected to exceed the rate of change of prices by approximately the amount of the increase in labor productivity, curve *CD* in Figure 16–4 is lower in relation to the vertical scale than curve *AB* in Figure 16–1 by 3 percent, which approximates the long-run average rate of increase in productivity in the American economy.

FIGURE 16-4
Modified Phillips Curve



A modified Phillips curve can easily be derived for the Perry study which was discussed earlier. Subtracting q from both sides of equation 14, we have

$$w - q = -6.814 + 23.240\frac{1}{U} - 1.580q + 0.670P \qquad (15)$$

Since, by Equation 12, $p = w - q$, we can substitute p in Equation 15 and write[23]

$$p = 6.814 + 23.240\frac{1}{U} - 1.580q + 0.670P \qquad (16)$$

Assuming, as we did earlier, that $q = 3$ and $P = 12$, Equation 16 reduces to

$$p = -3.514 + 23.240\frac{1}{U}$$

Based on this equation, the value of p associated with each value of U will be exactly three percentage points less than the corresponding value of w

[23] Perry, *op. cit.*, p. 61, Equation 3.15.

shown in Table 16–2. If the resulting values of p are plotted against values of U, they yield a modified Phillips curve very similar to *CD* in Figure 16–4.

According to curve *CD* in Figure 16–4, the general level of prices will be approximately stable when the unemployment rate is in the neighborhood of 6.5 percent (point P). In order to reduce unemployment to 4 percent of the labor force, it will be necessary to take measures to expand aggregate demand, and, by causing more liberal increases in money wage rates, this expansion of demand will lead to an increase in the general level of prices of about 2 percent per year (point Q). To reduce unemployment to 2.5 percent of the labor force by means of still more vigorous expansionary policies, it will be necessary to accept an increase in the price level of about 5 percent per year (point R).

The modified Phillips curve is designed to show the terms on which, by the use of general fiscal and monetary tools to control aggregate demand, the nation can trade reductions in unemployment for increases in the rate of inflation. The relationship is not an exact one but varies to some extent depending upon the attendant circumstances; nevertheless, our experience of recent years suggests that a relationship generally of this kind prevails in the American economy.

It is possible that by employing structural measures which will directly affect the functioning of labor markets—such as government intervention in the collective bargaining process in one form or another, measures to increase the mobility of labor, more vigorous enforcement of the antitrust laws to promote competition, and so on—the modified Phillips curve can be shifted to the left in such a way as to make a lower rate of unemployment compatible with price stability. For example, if the modified Phillips curve can be shifted by such measures from *CD* to *C′D′* in Figure 16–4, it will be possible to achieve an unemployment rate of about 4 percent under conditions of relative price stability (point S).[24]

Market-Power versus Demand-Pull Inflation

How can one recognize market-power inflation when it occurs and distinguish it from the demand-pull variety? The answer is that this is a distinction which is very difficult to make in practice. From mid-1946, when wartime price controls were repealed, to the end of 1969, the Wholesale Price Index rose by 90 percent. More than two thirds of this rise in prices is accounted for by two relatively brief flurries. The first extended from mid-1946 to the fall of 1948, when the index rose by nearly 50 percent, for

---

[24] The so-called "guideposts" for noninflationary wage and price behavior, first set forth in the *Economic Report of the President,* January 1962, pp. 185–90, were designed to influence the behavior of business and labor in such a way as to shift the modified Phillips curve to the left. For an evaluation of the guideposts, see John Sheehan, *The Wage-Price Guideposts* (Washington: Brookings Institution, 1967).

a rate of increase of more than 20 percent per year. The second occurred in the period from June 1950, to March 1951, when the index rose by 16 percent, again representing a rate of increase of over 20 percent per year. Most observers would agree that these two episodes of rapid inflation were caused by excessive aggregate demand—that of 1946–48 resulting from heavy pent-up demand for goods and services inherited from World War II, and that of 1950–51 being associated with the outbreak of the Korean War in June 1950. However, it would be difficult to prove conclusively that even these inflationary episodes were examples of pure demand-pull inflation.

Prices also rose enough to cause serious concern—although far less rapidly than in 1946–48 or 1950–51—during the period of so-called "creeping inflation" from the spring of 1955 to the spring of 1958, when the wholesale price index rose by nearly 9 percent, for a rate of increase of almost 3 percent per year; and as this is written (December 1969) prices have been rising since mid-1965 at a rate that has been a source of considerable concern.[25] While some observers would characterize the increases in prices in 1955–58 and in the current period as examples of market-power inflation, others would dispute this and attribute these increases to excessive aggregate demand.

Indeed, there are some economists who would deny the existence of market-power inflation as a distinguishable phenomenon, arguing that all inflation is necessarily a result of excessive aggregate demand. The basic argument of this school of thought is that even though there are cost pressures that at times tend to push up prices, the increase in prices cannot actually take place unless the inflationary tendencies are underwritten by a sufficient increase in aggregate demand to permit the output of goods and services to be purchased at the higher prices. While this argument is perfectly correct, the issue raised seems to be primarily semantic. It is both useful and, in principle at least, possible to distinguish between inflation *caused* by an excess of aggregate demand and inflation caused by autonomous upward adjustments of administered wages and prices which is then *underwritten* by an expansionary adaptation of aggregate demand.

Part of the difficulty in distinguishing between demand-pull and market-power inflation is that both types are associated with very similar observable

[25] The increases in the price level in the four inflationary periods, 1946–48, 1950–51, 1955–58, and 1965–69, add up to more than the overall increase of 90 percent for the entire period from mid-1946 to the end of 1969. The reason for this is, of course, that prices have fallen during some other portions of the period. The behavior of the Consumer Price Index has been generally rather similar to that of the Wholesale Price Index. However, the Consumer Price Index has risen almost continuously (though at varying rates) due primarily to the fact that the prices of many services, which are included in the Consumer but not in the Wholesale Price Index, have been rising steadily in the last few years. As indicated in Chapter 4, many observers believe that the Consumer Price Index has an especially serious upward bias, which causes it to overstate the rise in prices, because it does not make adequate allowance for improvements in the quality of goods and services.

characteristics. For example, money wage increases are almost certain to exceed increases in labor productivity in periods of inflation, whatever the origin of the inflation. In demand-pull inflation, excess demand will in due course make its appearance in labor markets, thereby pulling up money wages more rapidly than productivity rises. In market-power inflation, the initial impetus is very likely to come from upward adjustments in administered wages that exceed increases in productivity. It is true that the timing may be different: price increases may precede wage increases in demand-pull inflation, while the time pattern may be the other way around in the case of market-power inflation. However, these time patterns of adjustment of wages and prices are extremely difficult to trace through in practice; moreover, under certain patterns of business price determination, wage increases may precede price increases even in periods of demand-pull inflation.

One possibility that looks promising at first glance is to say that if inflation occurs at a time when the economy is operating below full employment, the inflation must be of the market-power variety. This method of making a distinction is certainly useful in extreme cases. If prices rise significantly when unemployment is 8 or 10 percent of the labor force, the causes of the inflation must almost certainly lie primarily in the exercise of market power. On the other hand, inflation that occurs when the unemployment rate is only 1 or 2 percent of the labor force (as in wartime) may be presumed to be primarily caused by excessive aggregate demand. Under more normal circumstances, however, this method of distinguishing the origins of inflation is not likely to be particularly useful. The difficulty lies in the fact that the concept of "full employment" is inherently fuzzy and to some extent arbitrary. A certain amount of "frictional" unemployment is normally associated with the process of economic change in a free economy —shifts in the pattern of demand for goods and services, changes in work preferences, movements of people from one part of the country to another, and so on. However, such frictional unemployment cannot be sharply distinguished from other kinds of unemployment, such as that due to a basic inadequacy of total demand. In conceptualizing "full employment," allowance must be made for frictional unemployment. Since the amount of frictional unemployment unquestionably varies with changing economic conditions, in ways that are hard to measure and define, the concept of full employment is practically impossible of precise analytical definition. And since, in the traditional view, demand-pull inflation occurs when aggregate demand exceeds the output that the economy can produce when it is operating at full employment, the inherent difficulty in defining precisely what constitutes full employment makes it impossible to specify in concrete terms the conditions under which pure demand-pull inflation is present.

Thus, under most conditions it is difficult to distinguish sharply between demand-pull and market-power inflation. Over the normal range within which economic activity fluctuates, prices tend to rise more rapidly as the unemployment rate is reduced. These movements of prices are the result of a complex interaction of cost adjustments and demand adjustments, which are extremely difficult to disentangle and identify. The relationship—as depicted by the modified Phillips curve in Figure 16–4—means that ordinarily, in order to reduce unemployment by expanding aggregate demand, we have to pay some price in the form of additional inflationary tendencies. Given the nature of this relationship, we, as a nation, have to decide, partly on economic grounds and partly in light of our general ethical preconceptions, what combination of unemployment and inflation is most acceptable to us. At the same time, we may also wish to consider the feasibility and weigh the possible benefits and costs of undertaking structural reforms which might shift the modified Phillips curve to the left, thereby enabling us to achieve more favorable combinations of unemployment and price stability.

PART IV

Economic Growth

INTRODUCTION

In the last decade, economic growth has been the subject of a great deal of discussion, not only among economists but also among government officials and the general public. During much of this period, growth has been considerably more rapid in the Soviet Union, Western Europe, and Japan than in the United States and this has been a cause of much concern here. Although the emphasis on growth has, at times, been somewhat exaggerated, the increased recognition of its importance has undoubtedly been healthy. Those responsible for economic policy have become fully aware of the necessity for demand to grow steadily if continuing full employment is to be achieved. And there has been increased recognition of the possibility that the government may by its policy actions be able to influence the rate at which the economy will grow under full employment conditions. In this part, we will set the analysis developed in previous chapters—in which growth was almost entirely neglected—in the more realistic context of a growing economy. We will also attempt to sort out the factors that determine the rate of growth and to consider whether—and how—it may be possible for the government to affect this rate.

Meaning and Measurement of Economic Growth

In an economy like that of the United States, in which the basic social objective is the well-being of the country's citizens, economic growth should be judged in terms of the progressive improvement of living standards. It is customary to use national income accounting concepts in measuring and evaluating growth. While these income concepts have some theoretical deficiencies as measures of growth, they are at least capable of fairly accurate quantitative estimation and therefore make growth a subject capable of concrete investigation.

Since consumption is the ultimate objective of economic activity in a nation with an individualistic orientation, personal consumption expenditures might seem to be an appropriate measure of economic performance and its rate of improvement. However, in an ongoing society, it is customary not to consume the entire current output of the economy but to set aside a portion of it by a process of saving and investment to make provision for the future. Consequently, consumption is too narrow a concept to use in measuring growth, since it does not include the entire output available but only that portion that the populace decides, individually and collectively, to use to satisfy immediate current needs.

At the other extreme, the most comprehensive measure of economic performance in the national accounts is the GNP. The most serious defect of the GNP as an index of well-being is that no deduction for the consumption of existing capital is made in computing it. GNP is a measure of the total output of the economy, including that portion of output used to make good the depreciation of existing capital. Thus, we could not, even in principle, consume the entire GNP year by year without progressively depleting the stock of capital.

In some respects, the best single concept to use in measuring economic performance and in studying economic growth would be the net national product (NNP). As was explained in an earlier chapter, the NNP is the GNP minus capital consumption allowances; the NNP is the total output of the economy, valued at market prices, after deduction for the consumption of existing capital.[1] Thus, it is a measure of the value of product that could, in principle, be consumed year by year without depleting the capital stock. However, there are some difficulties in using NNP. Economic performance and its rate of improvement require a measure that can be deflated—that is, expressed in dollars of constant value—since any rise in output or income that reflects merely the effect of rising prices does not add to human well-being. As we have seen in an earlier chapter, estimates of GNP valued at constant prices are regularly compiled by the Department of Commerce and are readily available in published sources on a current basis. Deflated values of NNP are difficult to prepare and are not compiled by the Department of Commerce. The reason is that it is difficult, both conceptually and practically, to prepare price indexes that can be used to deflate capital consumption allowances, primarily because such allowances represent mainly depreciation of capital goods that were purchased in past years at prices that have varied from year to year.[2] Deflated values for capital consumption allowances have been constructed for use in specialized

---

[1] See Chapter 2, p. 32.

[2] For further discussion of this point, see Chapter 2, pp. 32–33.

studies, but they are not available on a regular, current basis. Thus, lack of suitable deflated data limits the usefulness of NNP in the study of growth.

The ready availability of deflated data makes the GNP the most commonly used index of economic growth. There is, moreover, another reason for using GNP. As we will see, the maintenance of continuing growth under full employment conditions requires that aggregate demand expand at a rate sufficient to keep pace with the growth of productive capacity. As was pointed out in an earlier chapter, GNP is the appropriate income concept to use in the analysis of aggregate demand.[3] That is, the aggregate production which generates demand for labor—and puts pressure on the price level—includes the total production of capital goods, whether these goods are produced to add to the stock of capital (net investment) or to make good the depreciation of existing capital (replacement investment). Thus, in studying the balance between the growth of demand and the growth of productive capacity, GNP valued at constant prices is the best measure to use.

If real GNP is to be used as a measure of performance in the study of growth, should it be aggregate GNP or per capita GNP (that is, aggregate GNP divided by population)? Since the lot of the average person is improved only to the extent that the growth of aggregate GNP exceeds the growth of population, it would seem that per capita GNP is the preferred measure. However, if the growth of population is independent of the growth of GNP, it will not make much difference whether attention is focused on the growth of aggregate or per capita GNP. The assumption that the growth of population is independent of the growth of GNP is the one that is commonly made for working purposes; if the two growth rates are causally related, the links are complex and not well understood.[4] Moreover, to the extent that an analysis of aggregate demand is involved in the study of growth, it is more convenient to use aggregative rather than per capita GNP.

To summarize, on the basis of the considerations outlined above, it is common practice to use the percentage rate of growth of aggregate GNP valued at constant prices as a measure of economic growth. Accordingly, much of the analysis contained in this part will be conducted in these terms. Before proceeding with this analysis, however, it may be desirable to point out certain serious deficiencies of GNP—or, for that matter, any of the other readily available measures—as an index of growth.

1. In calculating GNP, as well as the other national income concepts, no adjustment is made for changes in disutility associated with the work effort

---

[3] See Chapter 3, pp. 61–62.

[4] There can be little doubt that the Great Depression of the 1930's markedly reduced marriage and birth rates, thereby slowing down population growth. In more normal times, however, it is doubtful whether moderate changes in the pace of economic activity have discernible effects on population growth.

expended in production. Thus, if the growth of GNP accelerates because workers are induced to expend more effort, some deduction should be made for the disutility of the additional effort in calculating the new growth rate, but the use of GNP as an index of growth will not involve such an adjustment. Conversely, if workers choose voluntarily to take some of the gains resulting from increased productive efficiency in the form of shorter working hours and more leisure time, the result will be a corresponding reduction in the rate of growth of GNP, even though the reduction in hours represents a rise in living standards for the workers. Failure to make any adjustment for the changes in the disutility of work effort is a serious fault with GNP as an index of growth, but a defect which is virtually impossible of correction.

2. For purposes of analyzing growth, the concept of investment and capital formation as incorporated in the national income accounts is more limited in scope than would be desirable. One difficulty is that consumer purchases of durable goods (except houses) such as automobiles, furniture, and household appliances are treated as current consumption in the period in which the goods are purchased.[5] A more suitable way of handling such purchases for purposes of analyzing growth would be to treat them as investment expenditures in the period in which they occur and to include in consumption and GNP an estimate of the imputed value of the flow of services emanating from the existing stock of such goods. As we will see, capital formation is an important source of economic growth, and the treatment of an important element of investment as though it were current consumption introduces distortions into the analysis of growth. A similar problem arises with respect to government investment. The construction of public projects such as schools, highways, public buildings, and recreation facilities should be treated as a form of investment in the period in which construction occurs, and the value of the services provided by the existing stock of such government capital facilities should be included in GNP. In practice, the concept of government capital formation is not recognized at all in the national income accounts; all government construction is simply treated as a current use of output.[6] Failure to handle consumer and government capital formation in a proper way in the national accounts is largely due to the practical difficulties of placing an appropriate valuation on the flows of services provided by these facilities. In addition, it can be argued that the treatment of these expenditures as though they were current consumption is more appropriate for the analysis of aggregate demand and employment, because consumer and government investment do not add to the productivity of the privately employed labor force in the same way as does investment by private business and therefore do not complicate the

---

[5] See Chapter 3, pp. 57–59.

[6] See Chapter 3, pp. 43–45.

problem of maintaining full employment.[7] But whatever the advantages of the present national income accounting methods for the purpose of analyzing aggregate demand, these methods are clearly not the ideal ones for the study of growth.

3. The GNP measure understates the rate of growth by an unmeasurable but substantial amount, because it excludes in large part the benefits that are derived from the introduction of new and improved products. Of course, all privately produced goods and services including new ones are included in the GNP at their current market value. However, as was pointed out in an earlier chapter, many of the price indexes currently in use contain an upward bias—that is, they generally tend to overstate price increases and understate price declines.[8] This bias, which is especially serious in the case of the Consumer Price Index, is primarily attributable to the virtual impossibility of taking account of the introduction of new products and of making adequate allowances for improvements in the quality of existing products. A similar bias exists in the case of price indexes of government services for which the prices are simply the wage rates of government employees, a practice which makes hours worked by such employees the direct measure of physical output, thereby precluding any adjustment at all for increases in productivity. The upward bias in the price indexes that are used as deflators in estimating GNP at constant prices produces a corresponding understatement of the growth of real output. This is illustrated by the following table:

	Year 1	*Year 2*	*Percent Change*
GNP (billions of current $)	600	655.2	9.2
"True" price deflator (year 1 = 100)	100	104.0	4.0
"Biased" price deflator (year 1 = 100)	100	105.0	5.0
GNP deflated by "true" price deflator	600	630.0	5.0
GNP deflated by "biased" price deflator	600	624.0	4.0

In this example, the biased price deflator shows an increase of 5 percent from year one to year two when the "true" price increase is 4 percent. But growth of GNP corresponding to the biased price deflator is 4 percent compared with the "true" increase of 5 percent. The failure of the GNP measure to capture the contribution of new consumer goods to the improvement of living standards is an extremely serious defect. As a result of the introduction of new surgical techniques and new medicines such as penicillin, man is now able to survive a variety of ills that formerly would have been fatal, and his life span has been greatly extended. Improved communi-

[7] W. L. Smith, "Consumer Instalment Credit: A Review Article," *American Economic Review,* Vol. XLVII, December 1957, pp. 966–84.

[8] See Chapter 4, pp. 85–88.

cations and transportation, brought about by the development of telephone, radio, television, automobile, and airplane, have vastly extended his horizons and the range of experience available to him. The great improvements in the human condition resulting from such major developments as these are not reflected—or at least show up only to a very minor extent—in the measured growth of real GNP.

The Significance of Economic Growth

Table A shows levels of aggregate GNP and per capita GNP that would be achieved in 1975, 1985, and 2000 for various growth rates starting from

TABLE A

Effects of Alternative Growth Rates on Aggregate and per Capita GNP, 1968–2000
(aggregate GNP, billions of 1968 dollars; per capita, 1968 dollars)

	1968 Actual	*Projections* 1975	1985	2000
		With 3.5% growth of aggregate GNP		
Aggregate GNP............	865.7	1,101.4	1,553.9	2,603/2
Per capita GNP............	4,303.0	5,022.0	6,144.0	8,457.0
		With 4% growth of aggregate GNP		
Aggregate GNP............	865.7	1,139.3	1,686.4	3,036.9
Per capita GNP............	4,303.0	5,195.0	6,668.0	9,866.0
		With 4.5% growth of aggregate GNP		
Aggregate GNP............	865.7	1,178.2	1,829.2	3,540.7
Per capita GNP............	4,303.0	5,373.0	7,233.0	11,503.0
		With 5% growth of aggregate GNP		
Aggregate GNP............	865.7	1,218.0	1,984.2	4,125.1
Per capita GNP............	4,303.0	5,554.0	7,846.0	13,402.0
Population (millions)*......	201.2	219.3	252.9	307.8

* Series C estimate of the Bureau of the Census.
SOURCES: *Survey of Current Business; Population Estimates*, Series P–25, February 20, 1967, "Projections of the Population of the United States by Age, Sex, and Color to 1990 with Extensions of Total Population to 2015," U.S. Department of Commerce, Bureau of the Census.

the actual GNP that prevailed in 1968. The calculations presented in this table demonstrate that even 3.5 percent growth of aggregate GNP can produce remarkable improvements in living standards in the course of a few years. Moreover, they show that a seemingly small increase in the growth rate of aggregate GNP can have an impressive effect on income per head of the population. For example, if GNP were to grow by 4 percent per year from 1968 to 1985, per capita GNP would be $6,668 in 1985, which would be $524, or 8.5 percent, greater than would be achieved in that year with a growth rate of 3.5 percent.

There is, however, another side to this kind of calculation. An increase of one half of a percentage point in the growth rate appears at first glance to be a small—indeed, almost a trivial—increase, and there is a tendency to think that such a tiny increase should be relatively easy to achieve. It should be noted, however, that an increase of one half point in the growth rate from 3.5 percent to 4 percent is a *relative* increase of one seventh, or about 14 percent. When the matter is looked at in this more appropriate way, it becomes apparent that an increase of one half of a percentage point is not so trivial after all. It is therefore perhaps less difficult to understand than might at first appear that, as we shall see, an increase of this magnitude in the growth rate may require considerable effort to achieve.

Chapter 17

MAINTAINING FULL EMPLOYMENT IN A GROWING ECONOMY

The analysis of the determination of the level of income that was developed in earlier chapters of this book was based on the assumption that the labor force and the stock of capital are constant and that there is no technological improvement. If this is the case, there is a certain level of aggregate real demand which, if maintained, will keep the economy operating at the desired level of employment. Of course, if the desired level of employment is associated with a progressive rise in the price level, along the lines developed in Chapter 16, aggregate demand in *monetary* terms will have to increase in pace with the price level. But a constant level of aggregate demand in *real* terms will suffice to maintain full employment.

These are not, however, the conditions that prevail in reality in a progressive, expanding economy such as that of the United States. The output that can be produced under conditions of full employment—however that term is defined—increases from month to month and year to year, although not necessarily at an entirely steady rate. For example, between 1948 and 1967, in both of which years unemployment was 3.8 percent of the civilian labor force, GNP valued at 1958 prices rose from $323.7 billion to $674.6 billion for an average increase of 3.9 percent per year. Part of the reason for this increase was growth of the civilian labor force from 60.6 million workers to 77.3 million. With the same unemployment rate, this average increase of 1.3 percent per year in the labor force brought a proportional increase in the number of workers employed. In addition, output per worker, valued at 1958 prices, rose from about $5,400 in 1948 to about $8,700 in 1967, an average increase of 2.5 percent per year. This increase in output per worker can be attributed in varying degrees to (*a*) technological advances that improved methods of producing goods and services, (*b*) accumulation of capital which provided the average worker with more and better tools and equipment to work with, and (*c*) improved education and training that raised the skill level of the average worker.

In an economy that grows in this way, it is clear that a constant level of aggregate demand will not suffice to maintain full employment, but rather that aggregate demand must expand at a sufficient rate to utilize the additional resources made available by growth. Thus, in order to be useful, the theory of income determination, which has been developed in the earlier chapters of this book, must be applied to the circumstances of a growing economy.

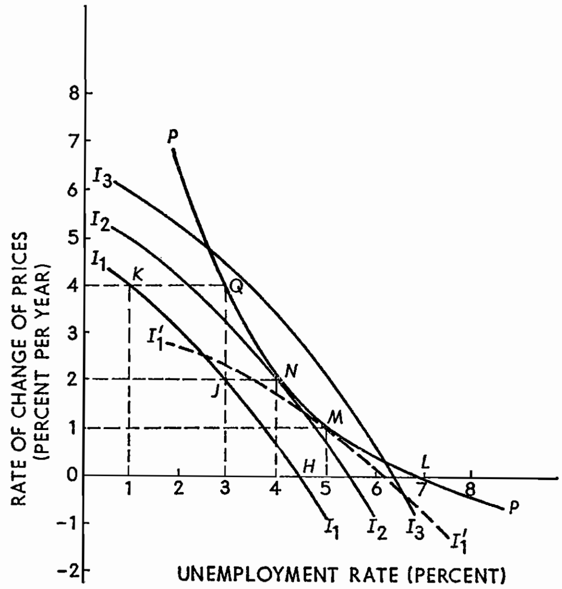
DEFINING FULL EMPLOYMENT

In an economy in which the rate of change of the price level is related to the rate of unemployment in accordance with the analysis presented in Chapter 16, it is necessary for the authorities responsible for economic policy to decide what their employment objective is—that is, to define what they mean by "full employment."

The kind of analysis that should, in principle, govern this decision is presented in Figure 17–1. The curve *PP* is the modified Phillips curve that

FIGURE 17–1

Modified Phillips Curve



was explained in Chapter 16.[1] It shows the "trade-off" that is assumed to exist between the rate of inflation (measured on the vertical axis) and the unemployment rate (measured on the horizontal axis). At an unemployment rate of 7 percent the price level will be stable (point *L*); at an unemployment rate of 5 percent, prices will rise at 1 percent per year (point

[1] See Chapter 16, pp. 366–68.

M); at 4 percent unemployment, prices will rise at 2 percent per year (point N); at 3 percent unemployment, prices will rise at 4 percent per year (point Q). The curves I_1I_1, I_2I_2, and I_3I_3, on the other hand, represent the preferences of those responsible for economic policy with respect to inflation and unemployment. Each of these curves is an indifference curve representing combinations of inflation and unemployment that are equally acceptable to the authorities. For example, the fact that points H, J, and K all lie on the same indifference curve (I_1I_1) indicates that the authorities would be equally willing to accept a stable price level combined with 4.5 percent unemployment (point H), 2 percent inflation combined with 3 percent unemployment (point J), or 4 percent inflation combined with 1 percent unemployment (point K). Curves that lie closer to the origin are preferred to those further out, for example, all points on I_1I_1 are preferred to all points on I_2I_2. This is apparent from the fact that each unemployment rate is associated with a lower rate of inflation on I_1I_1 than on I_2I_2. In establishing their preferences as reflected in the indifference curves, it may be supposed that the authorities are attempting to express the desires of the general public which they represent in the process of decision making.

In selecting the appropriate target with respect to inflation and unemployment, the authorities should choose from the economically possible combinations (those lying on the modified Phillips curve, PP), the one that lies on the lowest indifference curve. In this case, the optimal combination is represented by the point N—representing 4 percent unemployment and 2 percent inflation—since this point lies on I_2I_2, the lowest indifference curve touched by PP. That is, in this case we might define "full employment" as involving an unemployment rate of 4 percent, since this is the optimal employment target in light of the preferences of the authorities and the association that is judged to exist between inflation and unemployment.

It should be recognized that the above analysis makes the definition of full employment appear somewhat more scientific than it actually is. In reality, the authorities do not know exactly what the trade-off between inflation and unemployment—i.e., the modified Phillips curve, PP—actually is at any particular time. For example, in early 1961 when the unemployment rate was in the neighborhood of 7 percent, the incoming Kennedy administration had no way of knowing the precise way in which the price level would respond to a reduction in the unemployment rate. However, it formed a judgment that the rate of inflation that would be associated with an unemployment rate of 4 percent would not be intolerable, and it accordingly selected 4 percent as its employment target—that is, it defined full employment as 4 percent unemployment.[2] Nevertheless, Figure 17–1 represents a useful formalization of the rationale underlying the choice.

[2] It may be noted that the Kennedy administration referred to 4 percent unemployment as an "interim" target. That is, it hoped that once 4 percent had been achieved it would be possible by means of appropriate policies to reduce unemployment still further. For a discussion of the reasoning underlying the selection of the 4 percent target in 1961, see *Annual Report of the Council of Economic Advisers,* January 1962, pp. 44–48.

It is apparent from the analysis of Figure 17–1 that two administrations of different political complexion may define full employment differently even if they agree regarding the trade-off between inflation and unemployment (curve *PP*). The reason is that, reflecting the divergent views of their supporters, the two administrations may attach different relative degrees of importance to price stability as compared with high employment and, therefore, in terms of Figure 17–1, have different sets of indifference curves. Thus, for example, the optimal combination for one administration might be at point *N* (2 percent inflation and 4 percent unemployment) in Figure 17–1, as indicated above, whereas another administration might attach greater importance to price stability as reflected in indifference curve $I'_1I'_1$, and the optimum for it might be at *M* (1 percent inflation and 5 percent unemployment). Thus, for the first administration full employment would be 4 percent unemployment, whereas for the second, it would be 5 percent unemployment.

DEFINITION OF POTENTIAL OUTPUT OR PRODUCTIVE CAPACITY

Having chosen a definition of full employment along the lines developed above, it is useful to be able to translate this employment target into an associated target level of GNP. The GNP associated with full employment may be termed "potential output" or the "productive capacity" of the economy.

In 1961, Arthur Okun, then a member of the staff of the Council of Economic Advisers, devised a means of estimating the economy's potential output at an unemployment rate of 4 percent.[3] Based on the past relation between changes in real GNP and changes in the unemployment rate, and on movements of real GNP in relation to bench-mark periods when unemployment was in the neighborhood of 4 percent, Okun arrived at the following equation:

$$P = A[1 + 0.032(U - 4)] \tag{1}$$

where *P* is potential real GNP, or real GNP corresponding to a 4 percent unemployment rate; *A* is actual real GNP; and *U* is the unemployment rate (in percent). According to this equation, when the unemployment rate is 4 percent ($U = 4$), actual output is equal to potential output ($P = A$). This, of course, is a condition that would have to be met by a satisfactory estimating equation, since potential GNP is defined as the GNP that would prevail at a 4 percent unemployment rate.

[3] See A. M. Okun, "Potential GNP: Its Measurement and Significance," *1962 Proceedings of the Business and Economic Statistics Section of the American Statistical Association,* reprinted as *Cowles Foundation Paper No. 190* (Cowles Foundation for Research in Economics, Yale University, 1963), and reprinted in W. L. Smith and R. L. Teigen (eds.), *Readings in Money, National Income, and Stabilization Policy* (2d edition; Homewood, Ill.: Richard D. Irwin, Inc., 1970).

The estimates of potential GNP produced by the Okun equation should not be given an exact interpretation. For example, if the equation yields an estimate of $920 billion for a given period, this is best interpreted as meaning that potential GNP lies within a range of perhaps $5 billion around this figure—that is, between $915 billion and $925 billion. Moreover, since the equation is a purely empirical relation based on past behavior of the economy, it is likely to be less reliable as time passes and underlying conditions depart from those that prevailed in the period from which it is derived. But while it would be a mistake to take the equation too seriously, it does serve to illustrate certain basic relationships and ideas that are important.[4]

Perhaps the most interesting feature of the Okun analysis is its implication with respect to the relation between cyclical movements of output and employment. By simple algebraic manipulation, Equation 1 can be rewritten as follows:

$$\frac{P - A}{A} = 0.032(U - 4) \tag{2}$$

The expression $P - A$ is sometimes referred to as the "gap" between actual and potential GNP. Thus the left-hand side of Equation 2 is the gap taken as a percentage of actual GNP. The following table shows the size of the gap measured in this way for unemployment rates between 4 and 8 percent.

Unemployment Rate $(U)(\%)$	*Gap as % of Actual GNP* $[(P - A)/A]$
4	0.0
5	3.2
6	6.4
7	9.6
8	12.8

[4] For a more sophisticated model, incorporating some of the basic determinants of growth, which nevertheless yields estimates of potential output and the "gap" that are quite similar to Okun's, see R. M. Solow, "Technical Progress, Capital Formation, and Economic Growth," *American Economic Review*, Vol. LII, May 1962, pp. 76–86. It may be noted that in making estimates of the ratio of actual to potential output (A/P) in Table 2 (p. 82) based on Okun's model, Solow uses the equation

$$A = P[1 - 0.032(U - 4)]$$

This is different from the equation presented in Okun's article (*op. cit.*) and given in the text above, which is

$$P = A[1 + 0.032(U - 4)]$$

When this (presumably correct) form of Okun's equation is used, his results are even more similar to Solow's than Table 2 of the Solow article suggests. Another paper which uses a sophisticated model as the basis for estimating potential output is L. C. Thurow and L. D. Taylor, "The Interaction between the Actual and Potential Rates of Growth," *Review of Economics and Statistics*, Vol. XLVII, November 1966, pp. 351–60.

What this shows is that a reduction of one percentage point in the unemployment rate in the direction of 4 percent is associated with a 3.2 percentage point reduction in the gap between actual and potential output. To take a practical illustration, if potential output is $1,000 billion, actual output would be $912 billion at an unemployment rate of 7 percent[5] as compared with $1,000 billion at a 4 percent unemployment rate. Thus, a reduction of three percentage points in the unemployment rate would be associated with a 9.6 percent increase in actual GNP, from $912 billion to $1,000 billion.

The important point to note about this is the large gains in real output that are associated with rather modest reductions in unemployment. If output and employment bore a simple proportional relation to one another, an increase in the proportion of the labor force employed, from, say, 95 percent to 96 percent, would yield an increase of approximately 1 percent in total output. In fact, however, the Okun analysis indicates that a 1 percent increase in employment would raise total output by about 3 percent. Thus, reductions in unemployment within the range of unemployment rates from 4 to 8 percent not only alleviate the social and economic burden of unemployment on the unemployed themselves, but also yield magnified gains in real output for the economy in general.

There are several reasons for the fact that reductions in unemployment are associated with more than proportional increases in real output.

1. A reduction in unemployment tends to draw additional workers into the labor force. According to our statistics of employment and unemployment, a person is said to be in the labor force if he either has a job or is actively seeking work. When unemployment is high, some persons who would normally be seeking employment are discouraged from doing so and therefore leave the labor force. When unemployment declines and jobs become more readily available, these workers are attracted back into the labor force.[6] To illustrate, suppose that the labor force is initially 70 million

---

[5] This figure is obtained by solving the equation

$$1{,}000 = A[1 + 0.032(7 - 4)]$$

for A.

[6] Actually, two effects of changes in unemployment on labor force participation are present: a rise in unemployment may cause some workers to withdraw from the labor force because they become discouraged about the prospects of finding jobs; on the other hand, when the primary breadwinner in a family becomes unemployed, other members of the family may be induced to seek employment in an effort to maintain the family's income. The first of these effects has been termed the "discouraged worker effect," the second the "additional worker effect." A preponderance of evidence suggests that the "discouraged worker effect" is dominant, so that the labor force contracts as unemployment increases and expands as unemployment declines, as indicated in the text above; although there is some disagreement concerning the magnitude of the effect, which doubtless varies somewhat from one situation to another. For an excellent summary of the relevant research, see Jacob Mincer, "Labor-Force Participation and Unemployment: A Review of Recent Evidence," in R. A. Gordon and M. S. Gordon (eds.), *Prosperity and Unemployment* (New York: John Wiley & Sons, Inc., 1966), pp. 73–112.

workers, of whom 94 percent, or 65.8 million, are employed (i.e., the unemployment rate is 6 percent). An increase in demand which reduces the unemployment rate to 4 percent attracts an additional 0.5 million persons into the labor force, raising employment to 96 percent of 70.5 million, or 67.7 million, workers. Thus, employment is increased by about 1.9 million workers rather than only the 1.4 million that would have resulted from a decline of 2 percentage points in the unemployment rate with a constant labor force of 70 million. The induced expansion of the labor force has increased employment by half a million workers, and the productive contribution of these workers will expand the increment to output.

2. A reduction in unemployment tends to bring with it an increase in hours worked per man. That is, an increase in demand for labor is usually met in part by a lengthening of the work week. For this reason, also, the increase in labor input is greater than the fall in the unemployment rate would directly indicate.[7]

3. A decline in unemployment brings with it an increase in output per man-hour—that is, a rise in the average productivity of labor.[8] The main reason for this seems to be the existence of a substantial fixed element in labor cost. When demand falls, firms typically do not reduce their employment of sales and supervisory employees in proportion to the decline. They may also hesitate to lay off skilled workers because of the difficulty and cost of rehiring them again when demand expands and because of the effects on worker morale. Thus, when demand falls, there is an element of hidden unemployment in the form of employed workers who are less than fully utilized, and as a result output per worker declines. Conversely, when demand expands, the need for labor can partially be met by a fuller utilization of workers already on the payroll, and as a result output per worker increases. According to the classical view, one would expect productivity to decline during periods of rising employment and resource utilization, because less skilled workers would be added to payrolls and less efficient equipment would be drawn into use. While these effects are undoubtedly present, the statistical evidence drawn from the period since World War II clearly indicates that they are overbalanced by the more efficient utilization of overhead labor, with the result that productivity increases quite markedly as unemployment declines, at least within the moderate range of 4 to 8 percent unemployment rates.[9]

---

[7] On the relation between the unemployment rate and hours of work, see Okun, *op. cit.*, and Thurow and Taylor, *op. cit.*

[8] See Okun, *op. cit.*

[9] It should be noted that much of the fall in productivity that comes with an increase in unemployment is of a temporary or cyclical character. When aggregate demand falls and the decline is judged to be temporary, employers hesitate to lay off skilled workers, so that employment declines less than in proportion to the decline in production rates and labor productivity falls (or rises at less than its normal rate). When, in an ensuing period, demand again expands, it is possible to increase production rates with a less than

PROJECTING FUTURE GROWTH OF POTENTIAL GNP

The Okun equation was used for several years by the Council of Economic Advisers as a basis for making projections into the future of potential GNP, in order to be able to estimate the growth in aggregate real demand that would be needed to achieve or maintain an unemployment rate of 4 percent. Of course, the equation cannot be used directly for this purpose, because it is necessary to know actual GNP as well as the unemployment rate in order to use the equation to calculate potential GNP. However, Okun discovered that potential output, as calculated with his equation for the period 1954 to 1962, approximated quite closely to a simple 3.5 percent per year trend, starting from the actual GNP prevailing in mid-1955 when the unemployment rate was approximately 4 percent. Accordingly, the Council of Economic Advisers simply assumed that potential GNP was growing at 3.5 percent per year. One implication of such an assumption was that GNP must grow at about 3.5 percent per year in order to hold the unemployment rate constant. Later evidence, however, suggested that the growth of potential GNP had accelerated; for example, it appeared that GNP needed to grow more rapidly than 3.5 percent per year in order to prevent the unemployment rate from rising. Accordingly, the council increased its projected growth rate of potential GNP to 3.75 percent beginning with the first quarter of 1963,[10] and increased it again to 4 percent beginning with the first quarter of 1966.[11]

The Okun equation provides an estimate of what the GNP would be at 4 percent unemployment on the assumption that past relationships between output and unemployment continued to hold. Thus, it does not directly incorporate the underlying determinants of the growth of potential output or productive capacity, such as growth of the labor force, the rate of capital accumulation, and the pace of technological change. Accordingly, if any of these determinants change significantly, the estimate of potential output derived from the Okun equation may be quite wide of the mark. Moreover, there is no way to make an independent check of the accuracy of the estimates except to see what real GNP turns out to be when unemployment

normal increase in employment and productivity increases especially sharply. However, if the decline in aggregate demand is expected to last for a long time, employers will adjust their employment fully to the reduced rate of production. In such a situation, a rise in aggregate demand will not call forth a productivity bonus such as occurs during a recovery after a temporary or cyclical decline. Thus, if the economy has been operating for a long time at 6 percent unemployment and employers have fully adjusted their employment policies to this situation, an increase in aggregate demand that reduces unemployment to 4 percent will not result in a sharp rise in productivity such as would occur during a cyclical expansion that carried the economy from 6 percent unemployment to 4 percent unemployment.

[10] *Annual Report of the Council of Economic Advisers,* January 1965, p. 81.

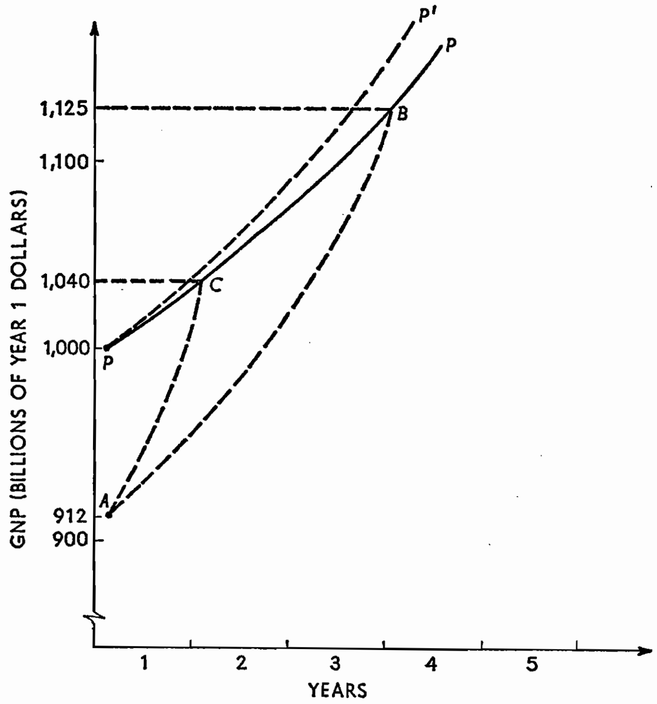
[11] *Annual Report of the Council of Economic Advisers,* January 1967, p. 44.

does reach the 4 percent level. It may be noted that when the unemployment rate fell to 4 percent in late 1965, actual GNP turned out to be very close to the Council of Economic Advisers' projection of potential GNP, which was based roughly on Okun's calculations.

FULL EMPLOYMENT AND ECONOMIC GROWTH

The relation between full employment and economic growth is illustrated in Figure 17–2. The solid line *PP* shows potential real GNP (defined

FIGURE 17–2
Relation between Full Employment and Economic Growth



as the GNP corresponding to 4 percent unemployment) growing at 4 percent per year, starting from a level of $1,000 billion (annual rate) in the first quarter of year one. Let us assume that actual real GNP in the first quarter of year one is $912 billion (point *A*), which corresponds to a 7 percent unemployment rate when potential GNP is $1,000 billion according to the Okun equation. If the authorities adopted expansionary monetary and fiscal policies to increase aggregate demand, actual GNP might expand along the dashed path *AB*, reaching the potential output path at point *B* in the first quarter of year four, at which time the unemployment rate would have been reduced to the desired level of 4 percent.

During this period of expansion toward full employment, growth would be very rapid—real GNP would rise from $912 billion in the first quarter of year one to $1,125 billion [$1,000 billion times $(1.04)^3$] in the first quarter of year four, for a rate of increase of 7.2 percent per year. However, as point *B* was approached, the rate of growth would have to be slowed down to 4 percent per year if actual GNP was to expand along the potential output or full employment path. If demand continued to rise more rapidly than 4 percent per year, unemployment would be reduced below 4 percent, presumably to the accompaniment of a greater rate of inflation than was judged acceptable.

When the economy is allowed to fall substantially below full employment to a position such as *A*, there is some kind of "speed limit" on the expansion back to full employment. It would be possible, in principle at least, to stimulate demand sufficiently by fiscal and monetary measures to restore full employment in the course of a year's time, expanding along a path such as *AC*. However, this would involve an increase in real GNP from $912 billion in the first quarter of year one, to $1,040 billion ($1,000 billion times 1.04) in the first quarter of year two, an increase of 14 percent in a single year. Such a rapid expansion would almost certainly lead to the appearance of bottlenecks in many product and factor markets as demand rose so rapidly that normal supply adjustments could not take place. Thus, some inflation of the bottleneck variety would occur, which could be avoided if the expansion to full employment were to occur at a more moderate pace. The "speed limit" problem is one that has not been subjected to any systematic research by economists, and the speed with which demand can be expanded to restore full employment without generating unnecessary bottleneck inflation is a matter concerning which judgments may differ. But there can be no doubt that when the economy has been permitted to fall substantially below full employment, to a point such as *A* in Figure 17–2, an effort to restore full employment too rapidly will generate inflation that could be avoided by a more moderate approach.

Full employment and a suitable rate of economic growth are often mentioned as separate goals of economic policy. In this context, growth means growth of potential output. Expansion toward full employment along a path such as *AB* in Figure 17–2 should not be viewed as growth. There are ways (to be discussed in the remaining chapters of this part of the book) in which economic policy might be able to tilt the growth curve of potential output upward—for example, to a path such as *PP′*, corresponding to a growth rate of 4.5 percent in Figure 17–2. It is best to reserve the concept of "growth policy" to refer to measures of this kind, and to regard measures to move the economy onto the growth path as "employment policy" or "stabilization policy."

In this chapter, we have for the most part viewed the growth of potential output as given and have paid almost no attention to its underlying determinants. This is a useful preliminary way to look at growth, but it is now time to depart from it and examine the determinants of growth. This we shall attempt to do in the next two chapters.

Chapter 18 CAPITAL ACCUMULATION AND ECONOMIC GROWTH: TWO SIMPLE MODELS

One of the major factors determining the growth of potential output or productive capacity is the rate of investment or capital accumulation. Net new investment adds to the stock of tools and equipment available for the labor force to work with; in addition, investment is the vehicle by which much new and improved technology is brought into effective use. In this section, we examine two rather simple models that have been employed by economists for the purpose of analyzing the relation of investment to growth.

THE DOMAR MODEL

As was pointed out earlier in this book, one of the shortcomings of Keynesian analysis is that it takes account of the effects of investment on aggregate demand but does not consider its impact on productive capacity.[1] Indeed, on the supply side of the Keynesian system, the only variable factor of production is labor. The capital stock is assumed to be fixed, and this assumption actually involves an inconsistency, since with net investment taking place the capital stock must in fact be growing. This inconsistency does not invalidate the usefulness of the Keynesian model for many purposes, but it does limit its value for the analysis of problems of economic growth.

One of the first attempts to take account of the effect of investment on productive capacity was made by Evsey D. Domar.[2] Domar assumed that net saving in year t is a constant fraction of net national product in year t—that is,

$$S_t = \alpha Y_t \tag{1}$$

[1] See Chapter 8, pp. 160–61.

[2] See E. D. Domar, "Expansion and Employment," *American Economic Review*, Vol. XXXVII, March 1947, pp. 34–55; and (for a mathematically more sophisticated presentation) "Capital Expansion, Rate of Growth, and Employment," *Econometrica*, Vol. XIV, April 1946, pp. 137–47. These papers are reprinted in Domar's *Essays in the Theory of Economic Growth* (New York: Oxford University Press, 1957), pp. 83–108 and 70–82, respectively.

where S is net saving, Y is net national product, and α is the marginal and average propensity to save. Then if we have

$$I_t = S_t \tag{2}$$

where I_t is net investment in year t, substituting Equation 1 into Equation 2 we obtain

$$I_t = \alpha Y_t$$

or

$$Y_t = \frac{1}{\alpha} I_t \tag{3}$$

Here $1/\alpha$ is the conventional Keynesian multiplier. Since no lags are allowed for, either in the adjustment of expenditures to income or in the adjustment of production to sales, it is assumed, in effect, that the multiplier works itself out fully within a year's time.[3]

The supply side of the economy is represented by the equation

$$\bar{Y}_t = \sigma K_t \tag{4}$$

where $\bar{Y}_t$ is full employment net national product in year t, K_t is the stock of capital at the beginning of year t and σ is the marginal (and average) output/capital ratio. That is, installation of an additional unit of capital will increase net output that can be produced at full employment by σ units. This same relation holds in period $t + 1$—that is,

$$\bar{Y}_{t+1} = \sigma K_{t+1} \tag{5}$$

Subtracting Equation 4 from Equation 5, we obtain

$$\bar{Y}_{t+1} - \bar{Y}_t = \sigma(K_{t+1} - K_t)$$

Since, by definition, $\Delta Y_t = Y_{t+1} - Y_t$ and $\Delta K_t = K_{t+1} - K_t$, we have

$$\Delta\bar{Y}_t = \sigma \Delta K_t$$

Moreover, since by definition $I_t = K_{t+1} - K_t = \Delta K_t$, we have

$$\Delta\bar{Y}_t = \sigma I_t \tag{6}$$

If full employment is achieved, Y_t must be equal to $\bar{Y}_t$, and this requires that there be the right amount of investment, so that, by Equation 3,

$$\bar{Y}_t = \frac{1}{\alpha} I_t \tag{7}$$

Solving Equation 6 for I_t, we have $I_t = \frac{1}{\sigma}\Delta\bar{Y}_t$. Substituting this for I_t in Equation 7, we obtain

[3] On the expenditure lag and the production lag, see Chapter 7.

$$\bar{Y}_t = \frac{1}{\alpha\sigma}\Delta\bar{Y}_t \qquad (8)$$

Rearranging Equation 8, we obtain

$$\frac{\Delta\bar{Y}_t}{\bar{Y}_t} = \alpha\sigma \qquad (9)$$

Since $I_t = \alpha\bar{Y}_t$ and $\Delta I_t = \alpha\Delta\bar{Y}_t$, from Equation 9 we also have

$$\frac{\Delta I_t}{I_t} = \alpha\sigma \qquad (10)$$

That is, the percentage rate of growth of income $(\Delta\bar{Y}_t/\bar{Y}_t)$ and the percentage rate of growth of investment $(\Delta I_t/I_t)$ are both equal to $\alpha\sigma$.

What the Domar model shows is that, on the quite restrictive assumptions (*a*) that net saving is a constant fraction, α, of net national product, and (*b*) that the marginal output/capital ratio, σ, is a constant, both net investment and net national product *must* grow at a constant rate of $\alpha\sigma$ *if full employment is to be maintained.* For example, if saving is 10 percent of income ($\alpha = 0.1$) and if an additional dollar of net investment increases net output at full employment by 30 cents ($\sigma = 0.3$), investment and income must grow at 3 percent per year in order to maintain full employment [$\alpha\sigma = (0.1)(0.3) = 0.03$]. A concrete numerical illustration for this case in which $\alpha = 0.1$ and $\sigma = 0.3$ may help to clarify the situation.

	Year		
	1	2	3
K	2,000	2,060.0	2,121.800
$\bar{Y}(= 0.3K)$	600	618.0	636.540
$Y(= C + I)$	600	618.0	636.540
$C(= 0.9Y)$	540	556.2	572.886
$I(= \Delta K)$	60	61.8	63.654
$S(= Y - C)$	60	61.8	63.654

At the beginning of year one, the stock of capital is 2,000 and, since $\sigma = 0.3$, full employment output is 600. If actual output is to be 600, net investment must be 60, sufficient to absorb the saving of 60 (10 percent of 600) that will occur at that level of income. If investment in year one is 60, this will increase the capital stock at the beginning of year two to 2,060, so the full employment output will rise to 618 (30 percent of 2,060) in year two. Thus, investment will have to rise to 61.8 (10 percent of 618) in that year if demand is to rise in pace with expanded productive capacity at full employment. This will raise the capital stock to 2,121.8 in year three, and so on. It is apparent that income rises from 600 in year one to 618 [(600)

(1.03)] in year two, to 636.54 [(600)(1.03)^2] in year three, etc.—for a rate of increase of 3 percent per year. Similarly investment rises at 3 percent per year.

It may be noted that if investment rises at a rate greater than $\alpha\sigma$, demand will increase more rapidly than capacity with inflationary consequences, while if investment increases at a rate less than $\alpha\sigma$, unemployment will appear. These two alternatives are illustrated in the following table:

		Year 2	
	Year 1	$\frac{\Delta I}{I} = 4\%$	$\frac{\Delta I}{I} = 2\%$
K	2,000	2,060.0	2,060.0
$\bar{Y}(= 0.3K)$	600	618.0	618.0
$Y(= C + I)$	600	624.0	612.0
$C(= 0.9Y)$	540	561.6	550.8
$I(= \Delta K)$	60	62.4	61.2
$S(= Y - C)$	60	62.4	61.2
Result	...	Inflation	Unemployment

It is assumed that the situation in year one is the same as in the earlier illustration and that full employment net national product is therefore 618 in year two. If instead of rising by 3 percent to 61.8 in year two investment rises by 4 percent to 62.4, demand for output will rise to 624, thereby exceeding full employment output, with inflationary consequences. On the other hand, if investment rises by only 2 percent to 61.2 in year two, net national output will be only 612. Since demand is less than the output that can be produced at full employment, there will be excessive unemployment.

It is important to recognize that the Domar model *contains no investment determination mechanism;* consequently, it says only that *investment must grow at the rate* $\alpha\sigma$ if continuous full employment is to be maintained. It does not address itself to the question of whether the necessary investment spending will be forthcoming and therefore whether full employment will in fact be maintained. In other words, Domar is concerned only with the investment requirements for continuous full employment.

In the Domar analysis, σ is the marginal productivity of capital, and the paper's conclusion is that in order to maintain full employment, the percentage increase in investment from one year to the next must be equal to the marginal productivity of capital times the saving rate. Moreover, it is assumed that the marginal productivity of capital is constant. However, the marginal productivity of capital, in general, is a function of the inputs of both capital and labor. Assuming that there is no technical change and that there are only two factors of production, the marginal productivity of capital is likely to remain approximately constant over time provided capital and labor are growing at the same rate.[4] In the above example, income (Y), investment (I), and the capital stock (K) all grow at the rate $\alpha\sigma$. If the labor force and employment also happen to be growing at the rate $\alpha\sigma$, then σ will probably remain roughly constant, and the analysis will hold. That is, in the above illustrative case, the analysis makes some sense if the labor force is growing at 3 percent per year. But, of course, in any particular case, it would be most unlikely that the labor force would be growing at the rate $\alpha\sigma$. Since the marginal productivity of capital (σ) is likely to fall if the ratio of labor to capital falls, σ will be falling rather than constant if the labor force is growing less rapidly than the stock of capital. With the saving rate constant, $\alpha\sigma$ will fall as σ falls—in other words, full employment can be maintained with investment increasing at a declining rate. However, if the labor force increases at a constant rate of n percent per year, an equilibrium may be reached when growth of capital relative to labor has reduced σ to the point at which $\alpha\sigma$ equals n, and from then on income, investment, the capital stock, and the labor force would all grow at the rate n. To illustrate, if the labor force was growing at 2 percent per year ($n = 0.02$) σ would gradually decline to an equilibrium value of 20 percent ($\sigma = n/\alpha = 0.02/0.1 = 0.2$). Conversely, if the labor force was initially growing at a rate greater than $\alpha\sigma$, σ would rise steadily until the point was reached where $\alpha\sigma$ became equal to n. Of course, the analysis would become much more complex in the (more realistic) case in which technical change was occurring.[5]

THE HARROD MODEL

Sir Roy Harrod has presented a model of economic growth which bears a superficial similarity to the Domar model but which in fact is somewhat different.[6] The main difference is that Harrod does have a theory of investment and attempts to explore the likelihood that growth may be self-sustaining.

---

[4] This will be true if the production function is characterized by constant returns to scale, that is, if an x percent increase in the inputs of both labor and capital will bring an x percent increase in output. See the discussion of returns to scale in Chapter 19, p. 402.

[5] Domar attempts to get around the difficulty referred to in this paragraph by defining σ ("Expansion and Employment," *op. cit.*, p. 90) as "the increase in productive capacity which *accompanies*, rather than which is caused by, each dollar invested" (italics in original), taking account of changes in the supplies of other factors and in technology. Thus, σ is distinguished from the marginal productivity of capital, which is the increase in output resulting from an additional dollar of capital invested, with quantities of other factors and the state of technology remaining constant. However, Domar's concept of σ causes difficulties because there is no reason why it should remain constant over time except under very special circumstances.

[6] See R. F. Harrod, *Towards a Dynamic Economics* (London: Macmillan & Co., Ltd., 1948), pp. 63–100; also his "An Essay in Dynamic Theory," *Economic Journal*, Vol. XLIX, March 1939, pp. 14–33.

Harrod assumes that there is a fixed proportional relation between output and the amount of capital needed to produce that output. Thus, we have

$$K_n = C_r Y_n$$

where Y_n is output (or net national product) in year n, K_n is the quantity of capital that must be in place at the beginning of year n if this output is to be produced, and C_r is the capital requirements coefficient.

Suppose we start at the beginning of the year t, and the stock of capital in existence at that time is K_t, and suppose further that businessmen in the aggregate expect to produce and sell enough output in year t to utilize fully the available capital. That is, expected output in year t (Y_t^e) satisfies the relation

$$Y_t^e = \frac{1}{C_r} K_t$$

Furthermore, businessmen look ahead to year $t + 1$ and plan their net investment in year t (I_t) so that they will have enough capital at the beginning of year $t + 1$ to be able to produce the amount they expect to sell in that year. This amount of capital is

$$K_{t+1} = C_r Y_{t+1}^e$$

where Y_{t+1}^e is expected sales and output in $t + 1$. By definition, we have

$$I_t = K_{t+1} - K_t$$

or

$$I_t = C_r Y_{t+1}^e - C_r Y_t^e = C_r (Y_{t+1}^e - Y_t^e) \tag{11}$$

Now suppose G is the percentage increase in sales and output that businessmen expect between year t and year $t + 1$; that is

$$Y_{t+1}^e = (1 + G) Y_t^e \tag{12}$$

Substituting Equation 12 into Equation 11, we obtain

$$I_t = C_r [(1 + G) Y_t^e - Y_t^e]$$

or

$$I_t = C_r G Y_t^e \tag{13}$$

Next let us suppose that saving (S) is a constant fraction (s) of income (Y); that is,

$$S_t = s Y_t \tag{14}$$

where Y_t is actual income in year t. If planned investment is adjusted to planned saving within a year's time (i.e., if there is no lag in the adjustment of production to sales), we have

$$I_t = S_t$$

or, using Equations 13 and 14,

$$C_r G Y_t^e = s Y_t \tag{15}$$

If businessmen's expectations about year t are to fulfilled, Y_t (actual income or production in year t) must equal Y_t^e (expected income or production in year t). Thinking of G as a variable, there is a certain value of G, which we will designate as G_w, that will make Y_t equal to Y_t^e. This value can be found by setting $Y_t^e = Y_t$ in Equation 15 and solving for G:

$$C_r G_w Y_t = s Y_t$$

or

$$G_w = \frac{s}{C_r} \tag{16}$$

G_w is what Harrod calls the *warranted rate of growth.* If s is 10 percent and C_r is 4, the value of G_w will be 2.5 percent ($0.1/4 = 0.025$). If businessmen estimate that output will grow at a rate of 2.5 percent per year and make their investment plans accordingly, their expectations for year t will be exactly fulfilled and the expansion will sustain itself at this rate.

A numerical illustration will help to clarify the Harrod model of growth. Let us assume that $C_r = 4$, $s = 0.1$, and the stock of capital (K_1) at the beginning of year one is 400. Then G_w ($= s/C_r$) will be 0.025 (or 2.5 percent per year). If businessmen do in fact expect sales and production to increase at the rate G_w, or 2.5 percent per year, the sequence of events will develop as in the table below. (The entries for employment [N_R and N_F] in the table will be discussed below.)

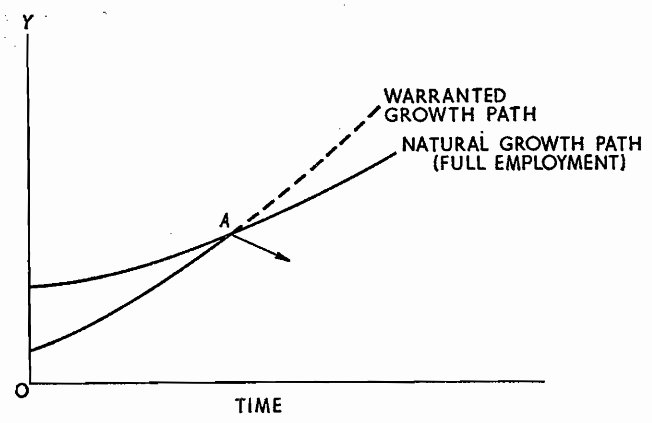
	Years		
	1	2	3
K	400	410.00	420.25000
Y^e	100	102.50	105.06250
$Y = C + I$	100	102.50	105.06250
$C = (1 - s)Y = 0.9Y$	90	92.25	94.55625
$I = C_r G_w Y^e = 0.1Y^e$	10	10.25	10.50625
$S = sY = 0.1Y$	10	10.25	10.50625
N_R	100	102.50	105.06250
N_F	102	103.02	104.05020
$N_F - N_R$	2	0.52	−1.01230

Expected income in year one (Y_1^e) is equal to $(1/C_r)K_1$, or 100. Output is expected to grow by 2.5 percent per year, so that $Y_2^e = 102.5$. Investment in year one is therefore $4(102.5 - 100) = 10$. Since the multiplier

$(1/s)$ is 10, this amount of investment will produce an income of 100 in year one which is equal to the income expected by businessmen. The capital stock in year two will be 410 $(= K_1 + I_1 = 400 + 10)$, and expected income in that year will be 102.5 $(= 1/4 \times 410)$. If income is again expected to increase by 2.5 percent between year one and year two, the reader can check for himself that actual income in year two will be 102.5, the same as expected income in that year. The entries for successive years can be traced out by the same process.

It is apparent that if businessmen happen to hit upon the *warranted rate of growth,* G_w $(= s/C_r)$, and base their investment plans on it, their expectations will be consistently fulfilled and steady growth at that rate will prevail. Note, however, that the system is unstable in the sense that if businessmen do not select exactly the right growth rate as the basis for their plans, the system will depart progressively farther from equilibrium. In the above example, if the rate of growth is expected to be 2 percent, expected income in year two will be 102 rather than 102.5, investment in year one will be 8 $[= 4\ (102 - 100)]$ rather than 10, and income in year one will be 80 $(= 10 \times 8)$ rather than 100. Thus, there will be unutilized capacity, which will cause a progressive decline in investment in succeeding periods. On the other hand, if a growth rate of 3 percent is selected, investment in period one will be 12, and total income 120. Since with the existing stock of capital only 100 can be produced, prices will have to rise, and an inflationary spiral will be set in motion.

In addition to the *warranted rate of growth* (G_w), Harrod refers to the *natural rate of growth* that would maintain full employment of labor. Harrod's model assumes fixed factor proportions. In the above table, the row labled N_R shows the amount of labor required to produce the indicated output—that is, 100 units of labor are required to produce 100 of output in year one, 102.5 units of labor are required to produce 102.5 of output in year two, and so on. Thus, in this example, the production of one unit of output requires four units of capital and one unit of labor. Unless the available labor force happens to grow at the rate G_w (and unless we start from a position of full employment), the warranted rate of growth does not coincide with the normal rate of growth, and growth at the warranted rate does not involve full employment of labor. In the above example, the available labor force (N_F) is 102 in year one and grows at a rate of 1 percent per year. For the first two periods, the labor force exceeds actual employment (N_R), and there is unemployment equal to the excess of N_F over N_R. However, by the third period a labor shortage is encountered because the warranted rate of growth of 2.5 percent per year exceeds the natural rate of growth of 1 percent per year. This labor shortage would, in fact, prevent the system from expanding at the rate G_w, and, in view of its inherent instability, this would send it into a tailspin. The result is shown conceptually in the following diagram.



Output "bounces off the full employment ceiling" at *A* and goes into a tailspin. Exactly how the economy would behave in these circumstances cannot be determined from the Harrod model, since the model only describes expansion along the warranted growth path.[7]

The Harrod model serves the useful purpose of illustrating in fairly simple terms the possibility of steady self-sustaining growth. However, it has several serious shortcomings. For one thing, its restrictive assumption of fixed capital and labor requirements, with no substitution possible between the two factors, gives the model an extreme rigidity and lack of plausibility. Another difficulty that renders the warranted rate of growth a dubious concept is the fact that investment decisions are made not by the economy as a whole but by individual firms, and not all of these firms will be expanding at the same rate—indeed, some firms may be declining even though the economy in general is growing.[8] Thus, while the model has some pedagogical value in illustrating the concept of self-sustaining growth, its practical significance is very doubtful.

The equation for the warranted rate of growth in the Harrod model (Equation 16) is $G_w = s/C_r$, while the conclusion to be drawn from the Domar model (Equation 9) is that in order for full employment to be maintained, income must grow at rate $\alpha\sigma$. Since s in Harrod and α in Domar both represent the average (and marginal) propensity to save, and since,

[7] Harrod's notion of the economy "bouncing off the ceiling" forms part of the basis of the theory of the business cycle developed by J. R. Hicks in his *A Contribution to the Theory of the Trade Cycle* (London: Oxford University Press, 1950).

[8] This is pointed out by T. C. Schelling in his "Capital Growth and Equilibrium," *American Economic Review*, Vol. XXVII, December 1947, pp. 864–76. Another objection raised by Schelling is that the fact that growth at the rate G_w utilized capacity fully in one period would not necessarily imply that producers would engage in a further equal *proportional* expansion in the next period as Harrod assumed. This same objection was raised by S. S. Alexander in "Mr. Harrod's Dynamic Model," *Economic Journal*, Vol. LX, December 1950, pp. 724–39.

in essence, σ ($\Delta Y/\Delta K$) in Domar is the reciprocal of C_r ($\Delta K/\Delta Y$) in Harrod, we have $\alpha = s$, and $\sigma = 1/C_r$, so that

$$(\alpha\sigma)_{\text{Domar}} = (s/C_r)_{\text{Harrod}}$$

The fact that the growth rates in the two models are the same—or at least very similar—has resulted in a tendency to lump the two together and refer to them jointly as the "Harrod-Domar model." The fact is, however, that, as explained above, there are substantial differences between the models.

CONCLUSION

We conclude this discussion by observing that while the Domar and Harrod models yield useful insights, they are too simplified to provide really satisfactory vehicles for analysis of the phenomenon of economic growth. Accordingly, in the next chapter we turn to more complex models that enable us to take account of factors other than capital formation that affect the rate of growth.

Chapter 19 MORE COMPLEX MODELS OF ECONOMIC GROWTH

One method of analyzing the determinants of productive capacity and its growth is to make use of a *production function.* The production function is one of the basic relationships of economic theory. It shows the output of a product as a function of the inputs of the factors of production used in producing it. For example, the production function of a firm might be written:

$$y = f(x_1, x_2, x_3, \ldots x_n)$$

where y is the output of the firm measured in physical units of product, and $x_1, x_2, \ldots x_n$ are the inputs of the n factors of production used by the firm also measured in physical units. The production function of a steel-producing firm, for instance, would show the number of tons of steel produced by the firm during a certain period of time as a function of the quantities of pig iron, limestone, coke, and other materials used, the number of man-hours of labor of various kinds employed, and the number of units of capital of various types used.[1]

There are certain properties that generally characterize production functions which should be set forth briefly. The first is the property of *diminishing returns* or *diminishing marginal productivity.* That is, if the quantities of all the factors except one are held constant, the additional output resulting from the application of an additional unit of that factor will decline as the quantity employed increases. Thus, with all factors other than labor held

[1] If, as is commonly the case, the firm produces several products, the production function might be

$$y_1 = f(x_1, x_2, \ldots x_n, y_2, y_3, \ldots y_m)$$

where y_1 is the quantity of product one produced, $x_1, x_2, \ldots x_n$ are the quantities of each of the n factors, and $y_2, y_3, \ldots y_m$ are the quantities of each of the other $m - 1$ products produced. That is, the quantity of product one produced is a function of the quantities of each of the factors employed and the quantities of each of the other products produced. This production function is sometimes written in implicit form

$$\phi(x_1, x_2, \ldots, x_n, y_1, y_2, \ldots y_m) = 0$$

The multiproduct case will not, however, concern us in the present context.

constant, if the addition of an 11th worker when 10 are already employed increases output by 10 units of product, then the addition of a 12th worker would increase output by less than 10, say 8, units of product. The explanation for diminishing returns is essentially that as more of one factor is employed with the quantities of other factors held constant, each unit of the variable factor has a reduced quantity of other factors to work with and output therefore does not rise in full proportion to the increased application of the variable factor.[2]

The second characteristic of production functions relates to *returns to scale*—that is, the way in which output changes as the *quantities of all factors are varied in the same proportion.* What will happen, for example, to output if the inputs of all factors are increased by 10 percent? There are three possibilities as follows:

Output Increases by	*Nature of Returns*
More than 10 percent	Increasing returns to scale
Exactly 10 percent	Constant returns to scale
Less than 10 percent	Decreasing returns to scale

If all factors are genuinely variable, constant returns to scale seems the most plausible result to expect under normal circumstances. If factors were combined in the optimal fashion initially, it should be possible to expand the scale of operations while maintaining this optimal combination.[3] However, it is possible in some cases that an increased scale of operations may permit the achievement of greater efficiency, thereby resulting in increasing returns to scale. Decreasing returns to scale would generally mean that some hidden factor of production, such as management, is not really being increased in proportion—that is, that the requirement of a proportional increase in all factors is not really being fulfilled.

AGGREGATE PRODUCTION FUNCTIONS

The concept of a production function can be applied to the economy as a whole—that is, the GNP can be related functionally to the quantities of capital and labor employed. Indeed, we made use of such an *aggregative production function* in a rudimentary way in our discussion of price flexibil-

[2] If the fixed factors are divisible so that the quantity actually used can be varied to match the inputs of the variable factor, it is possible to have constant marginal productivity over a certain range. However, this does not concern us in the present context.

[3] Since factors of production are available in discrete units, that is, are not infinitely divisible, it may not be possible to increase the inputs of all factors by a very small proportion—for example, a steel firm will not be able to increase the number of blast furnaces by 10 percent unless it already has 10 blast furnaces in operation. This means that there may be ranges within which production can only be increased by applying more variable factors to a certain quantity of fixed factors, thereby encountering diminishing returns. This point is, however, not relevant in the present context.

ity and employment in Chapter 15. The curve *OP* in Panel B of Figure 15–1[4] shows aggregate output (Y) as a function of the quantity of labor employed (N), and may be written algebraically as

$$Y = f(N, K)$$

where K is the stock of capital, which is assumed to be constant in accordance with the usual Keynesian analysis, and is therefore not depicted in the diagram. The fact that the curve *OP* becomes flatter as one moves out along it—i.e., that output does not increase in full proportion to the increase in employment—reflects the operation of diminishing returns as more labor is employed with a fixed amount of capital.[5] What we propose to do now is to extend this concept of the aggregate production function to consider circumstances in which the quantities of both labor and capital are variable, as a means of studying the determinants of growth.

Before proceeding with the discussion of aggregate production functions, however, it may be well to point out some pitfalls in this kind of analysis. It should be remembered that, as pointed out above, a production function is, strictly speaking, a relation between the *physical* output of a product and the *physical* inputs of various factors of production that are used to produce it. In the case of an *aggregate* production function, as usually formulated, output is represented by the GNP valued at constant prices, and inputs of labor and capital are represented by man-hours or man-years of labor employed and by an estimate of the dollar value of the aggregate capital stock deflated for price changes. This raises several difficulties.

1. In a genuine production function, labor of each different type and skill would be treated as a separate input, whereas in the aggregate production function all labor is treated as homogeneous and capable of being measured simply in aggregate man-hours. Thus, there can be important changes in labor input resulting from changes in the skill and occupational composition of the labor force that are not reflected in the measure of labor input used in an aggregate production function.

2. Both output (GNP) and the capital stock are measured in aggregate deflated dollar values and are therefore subject to changes of composition that are not reflected in the data. In truth there are many production functions in the economy—one for each producing firm—representing *physical* relationships between outputs and inputs, and the aggregate production function is a combination of these relationships weighted in a complex way

---

[4] See Chapter 15, p. 327.

[5] Curve *DD* in Panel A of Figure 15–1 also reflects diminishing returns to labor. The vertical height of this curve for any value of N corresponds to the slope of curve *OP* in Panel B for that same value of N. Thus, the height of *DD* in Panel A measures the marginal product of labor, and the fact that the curve has a negative slope is a reflection of diminishing marginal productivity.

by the relative values of output produced by these individual firms. Thus, an aggregate production function reflects not only the physical relationships between outputs and inputs of various firms but also the effects of changes in relative prices that occur as a result of changes in the supply of and demand for various products. That is, the aggregate production function reflects the entire operation of the economic system and not merely physical relationships between outputs and inputs.

3. In constructing aggregate production functions, inputs of materials are neglected on the ground that these materials are supplied by one firm and used by another and therefore cancel out for the economy as a whole, leaving only labor and capital as the basic factors of production. In fact, given the complexity of the relationships that exist among firms, this is not likely to be a legitimate assumption.

The student should not conclude, on the basis of these complications, that the aggregate production function is of no value in the study of the operation of the economy. However, in view of the fact that it is no more than a crude analogy to the legitimate concept of the production function that is employed in the theory of the firm, it is doubtful whether very fine-spun distinctions and theorems drawn from production theory have a useful role to play in the study of the economy as a whole.

The Cobb-Douglas Production Function

A particular form of production function that has been widely used and has proved to be helpful in the analysis of economic growth is the Cobb-Douglas production function which may be written as follows:[6]

$$Y = AK^{\alpha}N^{\beta}$$

where Y is output (GNP), K is the stock of capital, N is the quantity of labor employed, and A, α, and β are parameters. In general, α and β are positive and less than unity.[7]

---

[6] For a discussion of some of the earlier work involving the Cobb-Douglas production function, see P. H. Douglas, "Are There Laws of Production?" *American Economic Review,* Vol. XXXVIII, March 1948, pp. 1–41. The use of the Cobb-Douglas production function (as well as certain other types of production functions) in the analysis of economic growth is discussed in T. W. Swan, "Economic Growth and Capital Accumulation," *Economic Record,* Vol. XXXII, November 1956, pp. 334–61; and R. M. Solow, "A Contribution to the Theory of Economic Growth," *Quarterly Journal of Economics,* Vol. LXX, February 1956, pp. 65–94.

[7] The Cobb-Douglas production function can easily be written in logarithmic form. Taking logarithms of both sides of the equation, we have

$$\log Y = A' + \alpha \log K + \beta \log N$$

where $A' = \log A$. In this form, the equation is very easy to fit to data by regression techniques in order to obtain estimates of A, α, and β.

The nature of the returns to scale depends upon the value of $\alpha + \beta$, as follows:

$\alpha + \beta > 1$	Increasing returns to scale
$\alpha + \beta = 1$	Constant returns to scale
$\alpha + \beta < 1$	Decreasing returns to scale

In order to simplify matters somewhat, we shall assume constant returns to scale. In this case, $\beta = 1 - \alpha$, so we can write

$$Y = AK^{\alpha}N^{1-\alpha} \tag{1}$$

To illustrate the fact that this equation does show constant returns to scale, let us take a numerical example in which $A = 1$, and $\alpha = 1/2$. Then we have

$$Y = K^{\frac{1}{2}}N^{\frac{1}{2}}$$

Then if $K = \$400$ billion and $N = 100$ million man-years we have

$$Y = (400)^{\frac{1}{2}}(100)^{\frac{1}{2}} = \sqrt{400} \times \sqrt{100} = \sqrt{40{,}000}$$
$$Y = \$200 \text{ billion}$$

Suppose now that the inputs of capital and labor are both increased by 10 percent, to \$440 billion and 110 million man-years, respectively. Then we have

$$Y = (440)^{\frac{1}{2}}(110)^{\frac{1}{2}} = \sqrt{440} \times \sqrt{110} = \sqrt{48{,}400}$$
$$Y = \$220 \text{ billion}$$

Thus, a 10 percent increase in the inputs of capital and labor brings a 10 percent increase in output from \$200 billion to \$220 billion.

The parameter α is the elasticity of output with respect to capital input, and $1 - \alpha$ is the elasticity of output with respect to labor input. This is apparent from the following relation:[8]

$$\frac{dY}{Y} = \alpha\frac{dK}{K} + (1 - \alpha)\frac{dN}{N}$$

[8] This relation can be derived by the application of elementary calculus. Taking the total differential of Equation 1, we have

$$dY = \alpha AK^{\alpha-1}N^{1-\alpha}dK + (1 - \alpha)AK^{\alpha}N^{-\alpha}dN$$

Dividing this equation through by $Y = AK^{\alpha}N^{1-\alpha}$, we have

$$\frac{dY}{Y} = \frac{\alpha AK^{\alpha-1}N^{1-\alpha}dK}{AK^{\alpha}N^{1-\alpha}} + \frac{(1 - \alpha)AK^{\alpha}N^{-\alpha}dN}{AK^{\alpha}N^{1-\alpha}}$$

or

$$\frac{dY}{Y} = \alpha\frac{dK}{K} + (1 - \alpha)\frac{dN}{N}$$

If capital input increases with labor input constant ($dN = 0$), we have

$$\frac{dY}{Y} = \alpha \frac{dK}{K} \qquad (2)$$

That is,

Percentage increase in $Y = \alpha \times$ Percentage increase in K

Similarly, if labor input increases with capital input constant ($dK = 0$), we have

$$\frac{dY}{Y} = (1 - \alpha)\frac{dN}{N} \qquad (3)$$

or

Percentage increase in $Y = (1 - \alpha) \times$ Percentage increase in N

A numerical illustration will help to clarify these relationships. Suppose we have, as in our earlier example, $A = 1$, $\alpha = 1/2$, and initially, $K =$ \$400 billion, and $N = 100$ million man-years. Then

$$Y = (400)^{\frac{1}{2}}(100)^{\frac{1}{2}} = \$200 \text{ billion}$$

Now suppose capital input increases by 10 percent to \$440 billion with labor input unchanged. Then

$$Y = (440)^{\frac{1}{2}}(100)^{\frac{1}{2}} = \sqrt{440} \times \sqrt{100} = \sqrt{44{,}000}$$
$$Y = \$209.8 \text{ billion}$$

Thus, an increase of 10 percent in capital input brings an increase of 4.9 percent (from \$200 billion to \$209.8 billion) in output. That is, with $\alpha = 1/2$, the percentage increase in output is approximately one half of the percentage increase in capital input. The reason the increase in output is not precisely 5 percent is that the relationships in Equations 2 and 3 hold exactly only for infinitesimally small changes.

Technological Change

It is a relatively simple matter to introduce technological change into the Cobb-Douglas production function, at least in a limited way. This can be done by inserting a time trend as follows

$$Y = Ae^{rt}K^{\alpha}N^{1-\alpha} \qquad (4)$$

where e (2.71828 . . .) is the base of the Napierian system of logarithms, and r is the time rate of technological improvement.[9]

Several different kinds of technological change are generally recognized. Technological change is sometimes said to be *labor-saving* if it raises the marginal product of capital relative to the marginal product of labor, *capital-saving* if it raises the marginal product of labor relative to the marginal product of capital, and *neutral* if it raises the marginal products of labor and capital in the same proportion.[10] In this sense, the technological change we are allowing for is *neutral technological change.*[11] Another question is whether technological change must be *embodied* in capital equipment in order to become effective or whether it is *disembodied,* that is,

---

[9] The constant e can be derived as follows: suppose V is the value of \$1 invested at interest rate $100r$ percent for n years with interest compounded once a year—that is,

$$V = (1 + r)^n \qquad (i)$$

If interest is compounded twice a year, this becomes

$$V = \left(1 + \frac{r}{2}\right)^{2n}$$

and when interest is compounded m times a year, it becomes

$$V = \left(1 + \frac{r}{m}\right)^{mn} = \left(1 + \frac{r}{m}\right)^{\frac{mrn}{r}} = \left(\left(1 + \frac{1}{m/r}\right)^{m/r}\right)^{rn}$$

If we let $x = m/r$, the expression $[1 + 1/(m/r)]^{m/r}$ becomes

$$\left(1 + \frac{1}{x}\right)^x$$

It can be shown that as x becomes larger and larger—because compounding becomes more and more frequent—the value of $(1 + 1/x)^x$ approaches a certain number as a limit. This number is 2.71828 . . . $= e$. Thus, if compounding is continuous, Equation i above becomes

$$V = e^{rn}$$

[10] See J. R. Hicks, *The Theory of Wages* (London: Macmillan & Co., Ltd., 1932), pp. 121–27.

[11] Differentiating Equation 4 partially with respect to K and N, we obtain the marginal products of capital and labor,

$$\frac{\partial Y}{\partial K} = \alpha A e^{rt}\left(\frac{N}{K}\right)^{1-\alpha} = \text{marginal product of capital}$$

$$\frac{\partial Y}{\partial K} = (1 - \alpha)Ae^{rt}\left(\frac{K}{N}\right)^{\alpha} = \text{marginal product of labor}$$

Dividing $\frac{\partial Y}{\partial K}$ by $\frac{\partial Y}{\partial N}$, we obtain

$$\frac{\text{Marginal product of capital}}{\text{Marginal product of labor}} = \frac{\alpha}{1-\alpha}\frac{N}{K}$$

Since this expression does not involve the allowance for technological change, e^{rt}, it is apparent that the technological change is neutral in the sense indicated in the text.

where e (2.71828 . . .) is the base of the Napierian system of logarithms, and r is the time rate of technological improvement.[9]

Several different kinds of technological change are generally recognized. Technological change is sometimes said to be *labor-saving* if it raises the marginal product of capital relative to the marginal product of labor, *capital-saving* if it raises the marginal product of labor relative to the marginal product of capital, and *neutral* if it raises the marginal products of labor and capital in the same proportion.[10] In this sense, the technological change we are allowing for is *neutral technological change.*[11] Another question is whether technological change must be *embodied* in capital equipment in order to become effective or whether it is *disembodied,* that is,

---

[9] The constant e can be derived as follows: suppose V is the value of $1 invested at interest rate $100r$ percent for n years with interest compounded once a year—that is,

$$V = (1 + r)^n \qquad (i)$$

If interest is compounded twice a year, this becomes

$$V = \left(1 + \frac{r}{2}\right)^{2n}$$

and when interest is compounded m times a year, it becomes

$$V = \left(1 + \frac{r}{m}\right)^{mn} = \left(1 + \frac{r}{m}\right)^{\frac{mrn}{r}} = \left(\left(1 + \frac{1}{m/r}\right)^{m/r}\right)^{rn}$$

If we let $x = m/r$, the expression $[1 + 1/(m/r)]^{m/r}$ becomes

$$\left(1 + \frac{1}{x}\right)^x$$

It can be shown that as x becomes larger and larger—because compounding becomes more and more frequent—the value of $(1 + 1/x)^x$ approaches a certain number as a limit. This number is 2.71828 . . . $= e$. Thus, if compounding is continuous, Equation i above becomes

$$V = e^{rn}$$

[10] See J. R. Hicks, *The Theory of Wages* (London: Macmillan & Co., Ltd., 1932), pp. 121–27.

[11] Differentiating Equation 4 partially with respect to K and N, we obtain the marginal products of capital and labor,

$$\frac{\partial Y}{\partial K} = \alpha A e^{rt} \left(\frac{N}{K}\right)^{1-\alpha} = \text{marginal product of capital}$$

$$\frac{\partial Y}{\partial K} = (1 - \alpha) A e^{rt} \left(\frac{K}{N}\right)^{\alpha} = \text{marginal product of labor}$$

Dividing $\frac{\partial Y}{\partial K}$ by $\frac{\partial Y}{\partial N}$, we obtain

$$\frac{\text{Marginal product of capital}}{\text{Marginal product of labor}} = \frac{\alpha}{1 - \alpha} \frac{N}{K}$$

Since this expression does not involve the allowance for technological change, e^{rt}, it is apparent that the technological change is neutral in the sense indicated in the text.

exists as an independent force affecting output without requiring new investment in equipment to become operative. The issue of embodied versus disembodied technological change has important implications for economic growth, which we will discuss below. It should be noted at this point that the technological change allowed for in Equation 4 is of the disembodied variety; that is, it operates independently of investment and does not need to be incorporated in capital equipment to become effective.

In Equation 4, the factors affecting output are broken into three categories: technical change (e^{rt}), capital (K), and labor (N). In this context, technical change necessarily takes on a broad meaning—in effect, it is a residual which accounts for everything affecting output other than the aggregate private capital stock and the aggregate supply of labor. Thus, it includes not only the effects of inventions and technical innovations but also the effects of such other developments as (*a*) increases in labor skills resulting from education and training, and (*b*) public investment in social overhead capital, such as highways, to the extent such investment indirectly increases the efficiency of private industry.

The following expression involving rates of growth can be derived from Equation 4[12]

$$y = r + \alpha k + (1 - \alpha)n \tag{5}$$

where y is the percentage rate of growth of output (GNP), r is the percentage rate of technological improvement, k is the percentage rate of growth of the capital stock, and n is the percentage rate of growth of the supply of labor. That is, the rate of growth of output may be accounted for as the sum of three components:

1. The rate of technological improvement (r).
2. The elasticity of output with respect to the capital stock (α) times the rate of growth of the capital stock (k).
3. The elasticity of output with respect to labor input ($1 - \alpha$) times the rate of growth of labor input (n).

---

[12] Differentiating Equation 4 totally with respect to time yields

$$\frac{dY}{dt} = rAe^{rt}K^{\alpha}N^{1-\alpha} + \alpha Ae^{rt}K^{\alpha-1}N^{1-\alpha}\frac{dK}{dt} + (1 - \alpha)Ae^{rt}K^{\alpha}N^{-\alpha}\frac{dN}{dt}$$

Dividing through by $Y = Ae^{rt}K^{\alpha}N^{1-\alpha}$, we obtain

$$\frac{1}{Y}\frac{dY}{dt} = r + \alpha\frac{1}{K}\frac{dK}{dt} + (1 - \alpha)\frac{1}{N}\frac{dN}{dt}$$

Letting $y = (1/Y)(dY/dt)$ = rate of growth of output, $k = (1/K)(dK/dt)$ = rate of growth of capital, and $n = (1/N)(dN/dt)$ = rate of growth of labor, we obtain Equation 5 in the text. It may be noted that y, k, n, and r are instantaneous percentage rates of growth, that is, they involve continuous compounding as explained in footnote 9. However Equation 5 holds approximately for ordinary annual percentage changes.

For example, if technical progress is proceeding at a rate of 1 percent per year, the capital stock is growing at 2 percent per year, the labor force is growing at 1.5 percent per year, and $\alpha = 0.3$, the rate of growth will be

$$y = 1 + (0.3)(2) + (0.7)(1.5) = 2.65 \text{ percent per year}$$

Now let us suppose that net saving (S) under conditions of full employment is a constant proportion (s) of GNP; that is,

$$S = sY$$

If full employment is to be maintained it will be necessary for net investment (I) to be kept equal to the net saving that will occur at full employment. This would generally require that appropriate fiscal and monetary policies be adopted. In this context, net saving is, of course, not merely personal saving; rather it includes personal saving plus corporate saving (corporate retained earnings) plus the surplus (or minus the deficit) in the government budget.

No matter what the initial capital endowment of the economy is in relation to income, if the saving rate, s, is established at a certain level and remains there, the rate of growth of capital will gradually approach equality with the rate of growth of income—that is, y will approach equality with k.[13] Thus, the equilibrium rate of growth for any given saving rate can be obtained by setting k equal to y in Equation 5 above, which yields for the equilibrium rate of growth of income

$$y = n + \frac{r}{1 - \alpha}$$

Similarly, the equilibrium rate of growth of capital is

$$k = n + \frac{r}{1 - \alpha} \tag{6}$$

The interesting point about this result is that it shows that the equilibrium rate of growth is independent of the saving rate, s. That is, an increase in the saving rate does not permanently raise the rate of growth.

It should be noted, however, that investment is the increment to the capital stock, so that the percentage growth of capital is equal to I/K; moreover $I = S = sY$. Thus

$$k = \frac{I}{K} = \frac{sY}{K}$$

[13] This is intuitively plausible: if income is growing and the increments to capital resulting from saving and investment are a constant fraction of income, it seems reasonable that the rate of growth of the stock of capital will gradually approach the rate of growth of income. It can in fact be rigorously proved that, for the Cobb-Douglas model with neutral technological change and a constant saving rate, k will indeed approach y asymptotically.

Substituting this in Equation 6, we have, in equilibrium,

$$\frac{sY}{K} = n + \frac{r}{1 - \alpha}$$

or

$$\frac{K}{Y} = \frac{s(1 - \alpha)}{n(1 - \alpha) + r} \tag{7}$$

That is, *even though an increase in the saving rate does not permanently increase the rate of growth, it does raise the equilibrium capital-output ratio in direct proportion.*

A numerical illustration may help to clarify these relationships. Suppose the labor supply grows at 1.5 percent ($n = 0.015$), the rate of technological change is 1.4 percent ($r = 0.014$), the elasticity of output with respect to capital input is 0.3 ($\alpha = 0.3$), the elasticity of output with respect to labor input is 0.7 ($1 - \alpha = 0.7$), and the labor supply (N) is initially 437 million units. Under these conditions, the equilibrium rate of growth of income and capital is

$$y = k = n + \frac{r}{1 - \alpha} = 0.035 \text{ or } 3.5 \text{ percent}$$

To begin with, let us suppose that the saving rate is 10.5 percent ($s = 0.105$); this yields an equilibrium capital-output ratio of

$$\frac{K}{Y} = \frac{s(1 - \alpha)}{n(1 - \alpha) + r} = 3$$

If we assume that the economy is initially (i.e., when $t = 0$) on the equilibrium growth path corresponding to the given labor force and capital-output ratio, then GNP at that time would be \$700 billion (and the capital stock would have to be \$2,100 billion).[14] If the saving rate remains at 10.5 percent and if necessary measures are taken to insure that all of the savings

---

[14] Designating the initial labor supply as N_0, the labor supply in any period t will be N_0e^{nt}. The equilibrium stock of capital is given by $K = [s(1 - \quad)/(n(1 - \alpha) + r)]$, from Equation 7. Substituting these values into Equation 4, we have

$$Y_t = A_e^{rt}\left(\frac{s(1 - \alpha)}{n(1 - \alpha) + r} Y_t\right)^{\alpha} (N_0e^{nt})^{1-\alpha}$$

or

$$Y_t = A\left(\frac{s(1 - \alpha)}{n(1 - \alpha) + r}\right)^{\alpha} Y_t^{\alpha} N_0^{1-\alpha} e^{[r+n(1-\alpha)]t}$$

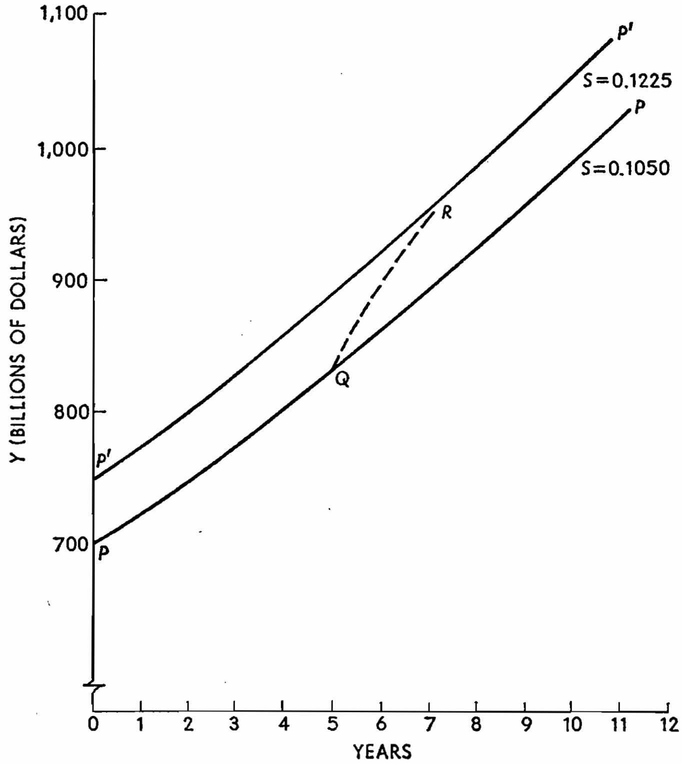
Solving explicitly for Y_t (which appears on both sides of this equation), we obtain

$$Y_t = \left[A\left(\frac{s(1 - \alpha)}{n(1 - \alpha) + r}\right)^{\alpha}\right]^{\frac{1}{1-\alpha}} N_0 e^{(n+r/1-\alpha)t}$$

Assuming for simplicity that $A = 1$ and substituting the values given above ($s = 0.105$,

FIGURE 19–1

Illustrative Growth Paths



generated by this saving rate at full employment are invested, the economy will grow along the path *PP* in Figure 19–1, starting with a GNP of $700 billion when $t = 0$ and growing at a rate of 3.5 percent.

Now suppose that at the end of the fifth year, the saving rate rises to 12.25 percent ($s = 0.1225$). The equilibrium growth rate is still 3.5 percent, since, as explained earlier, this rate is independent of the saving rate. However, the equilibrium capital-output ratio, K/Y, rises to 3.5. The equilibrium growth path corresponding to a saving rate of 12.25 percent is represented by *P′P′* in Figure 19–1; this path is above the path *PP* corre-

---

$\alpha = 0.3$, $n = 0.015$, $r = 0.014$, and $N_0 = 437$), we obtain

$$Y_t = (3)^{3/7}(437)e^{0.035t}$$

For $t = 0$, we have (since $e^0 = 1$)

$$Y_0 = (3)^{3/7}(437) = \$700 \text{ billion}$$

The growth path *PP* corresponding to a 10.5 percent saving rate in Figure 19–1 then becomes simply

$$Y_t = Y_0e^{0.035t} = 700e^{0.035t}$$

sponding to a saving rate of 10.5 percent, but the rate of growth is the same, 3.5 percent.[15] When the saving rate rises, the rate of growth will increase (assuming, of course, that measures are taken to insure that the additional saving is matched by investment), and GNP will move along the dashed path *QR* in Figure 19–1. The capital-output ratio will gradually increase from its original equilibrium value of 3, corresponding to path *PP,* and a saving rate of 10.5 percent toward its new equilibrium value of 3.5, corresponding to path *P′P′* and a saving rate of 12.25 percent. As time passes, the rate of growth will slow down toward its original value of 3.5 percent and the economy will move gradually toward the higher growth path *P′P′*. In a sense, the rise in the saving rate will not permanently affect the growth rate, but it will nevertheless permanently raise GNP above the levels that would have prevailed had the saving rate remained at 10.5 percent.

Thus, if the saving rate is constant, the rate of growth of income—and also the rate of capital accumulation—is ultimately determined by the rate of growth of labor input (n), the rate of technical progress (r), and the elasticity of output with respect to capital (α), in accordance with the equation:

$$y = k = n + \frac{r}{1 - \alpha}$$

Any increase in n, r, or α will, therefore, raise the growth rate.

It should be emphasized that the above analysis relates to the growth of potential or capacity output under conditions of full employment, as defined in accordance with the analysis presented in Chapter 17, and is predicated on the assumption that fiscal and monetary policies are adopted which are successful in keeping the economy operating continuously at full employment.[16] In practice, the inducement to invest is likely to undergo fluctuations reflecting, for example, the operation of the acceleration principle as explained in Chapter 9 and this introduces complications into the analysis of economic growth.

1. If, as is very likely, it does not prove possible to make adjustments in policy that fully compensate for fluctuations in investment demand, GNP

---

[15] The equation of the growth path *P′P′*, corresponding to that given in footnote 14 for path *PP* is

$$Y_t = (3.5)^{3/7}(437)e^{0.035t}$$

For $t = 0$

$$Y_0 = (3.5)^{3/7}(437) = \$748 \text{ billion}$$

Thus, the growth path becomes simply

$$Y_t = Y_0e^{0.035t} = 748e^{0.035t}$$

[16] If prices and wages were fully flexible, a growing economy might automatically maintain itself at full employment in accordance with the analysis presented in Chapter 15, pp. 325–29. However, the necessary degree of price and wage flexibility is clearly not present in reality, and appropriate adjustments in fiscal and monetary policy would therefore in fact be necessary.

may undulate around the full employment growth path, dropping below the path in periods of less than full employment and rising above it in periods of inflation and over-full employment. It may be noted that these fluctuations in actual GNP will be reflected in cyclical movements of the growth of potential or capacity output. Declines in investment in recession periods will reduce the rate of capital accumulation and slow down the growth of potential output, and unusually high levels of investment in periods of prosperity will accelerate the pace of capital accumulation and cause potential output to grow more rapidly.

2. Even if it is possible by means of fiscal policy to offset perfectly the effects on aggregate demand of fluctuations in the inducement to invest and thereby keep the economy operating continuously at full employment, there will nevertheless be fluctuations in the growth of potential and actual output. These fluctuations will show up in the form of variations in the net saving rate. When investment demand declines, it will be necessary to reduce taxes or increase government expenditures, thereby moving the government budget in the direction of a deficit. Since the government surplus is a portion of the net saving of the economy, such a movement toward a deficit will constitute a reduction in the saving rate. Conversely, when investment demand is unusually strong, taxes will need to be increased or government expenditures reduced to prevent aggregate demand from rising above full employment levels. This will move the budget toward a surplus and increase the net saving rate. Thus, compensating changes in saving and investment will cause capital accumulation and the rate of growth of output to undergo fluctuations.

3. If the economy is allowed to fall substantially below full employment and to remain in such a condition for a prolonged period of time, the growth of potential output may be sharply reduced as the inducement to invest is weakened. Historically, the most serious case of this kind is the 1930's, when the Great Depression weakened investment incentives so severely that productive capacity grew hardly at all for a whole decade. A similar but less severe case was the period from 1957 to perhaps 1963, when unemployment fluctuated between 5 and 7 percent and there was a sufficient loss of confidence on the part of businessmen to keep investment at abnormally low levels. It may be noted that a prolonged period of unemployment and weak markets for output may not only reduce investment but also induce a cautious and unprogressive attitude on the part of business and labor that is not conducive to the introduction of those technological improvements that are vital to rapid economic growth.

INVESTMENT VERSUS TECHNOLOGICAL CHANGE AS SOURCES OF ECONOMIC GROWTH

In terms of the framework developed above, the rate of growth of productive capacity depends upon the rate of growth of the full employment

or

$$y - n = r + \alpha(k - n)$$

Now $y - n$ is the rate of growth of output per man-hour, and $k - n$ is the rate of growth of capital per man-hour. Thus, if output grows at 3 percent per year while the number of man-hours grows at 2 percent per year, output grows 1 percent faster than man-hours, or output per man-hour grows at a rate of 1 percent per year. If we let $y_n = y - n =$ the rate of growth of output per man-hour and $k_n = k - n =$ the rate of growth of capital per man-hour, we have

$$y_n = r + \alpha k_n \qquad (8)$$

It is an elementary proposition in the theory of the firm that under conditions of pure competition—that is, when the individual firm takes the price it receives for its product and the prices it must pay for factors of production as given and not capable of being influenced by its own actions—each factor of production will be hired up to the point at which its price is equal to the value of its marginal product. The value of the marginal product (VMP), in turn, is equal to the marginal physical product of the factor (MPP) multiplied by the price of the product (P). That is

$$VMP = MPP \times P$$

Thus, we have

$$Q = VMP = MPP \times P$$

where Q is the price of the factor. The real remuneration of the factor is therefore

$$\frac{Q}{P} = MPP$$

That is, under conditions of pure competition, each factor receives a real rate of remuneration equal to its marginal physical product. Thus, if there are only two factors of production, capital (K) and labor (N), the real remuneration of a unit of capital is equal to the marginal physical product of capital (MPP_K) and the real remuneration of a unit of labor is equal to the marginal physical product of labor (MPP_N). The aggregate real income of each factor is obtained by multiplying the quantity of that factor in use by its remuneration per unit; that is

$K \times MPP_K =$ Aggregate real income received by owners of capital
$N \times MPP_N =$ Aggregate real income received by labor

The proportional shares of total income (Y) going to capital and labor are

$\frac{K \times MPP_K}{Y}$ = Fraction of total income received by the owners of capital.

$\frac{N \times MPP_N}{Y}$ = Fraction of total income received by labor

If capital and labor are the only factors of production, these two fractions must account for total income—that is, they must add up to unity:

$$\frac{K \times MPP_K}{Y} + \frac{N \times MPP_N}{Y} = 1$$

It can be shown for the Cobb-Douglas production function (Equation 4) that under conditions of competition in which capital and labor receive their respective marginal products, α is the share of income received by capital.[18] That is,

$$\alpha = \frac{K \times MPP_K}{Y}$$

Solving Equation 8 for r we have

$$r = y_n - \alpha k_n \tag{9}$$

If we now think of r, y_n, α, and k_n as varying from year to year rather than remaining constant, values of y, α, and k can be calculated, and Equation 9 can be used to make an estimate of r. This can be illustrated by Solow's calculation for 1922:[19]

	Private Nonfarm GNP per Man-Hour	*Employed Capital per Man-Hour*
1922	$0.788	$2.49
1923	$0.809	$2.61
Increase	2.7%	4.8%

[18] Using the expression for the marginal product of capital derived in footnote 11, the share of income going to capital is

$$\frac{K \times MPP_K}{Y} = \frac{K\frac{\partial Y}{\partial K}}{Y} = \frac{K\alpha Ae^{rt}\left(\frac{N}{K}\right)^{1-\alpha}}{Ae^{rt}K^{\alpha}N^{1-\alpha}} = \frac{\alpha Ae^{rt}K^{\alpha}N^{1-\alpha}}{Ae^{rt}K^{\alpha}N^{1-\alpha}} = \alpha$$

The share of income going to labor can be shown similarly to be equal to $1 - \alpha$.

[19] Government and agriculture were excluded from Solow's calculations: thus, he used private nonfarm GNP rather than total GNP. Since the unemployment rate varied considerably from year to year during the period covered and since the portion of the capital stock actually in use presumably varies with employment of labor, Solow made a crude adjustment to allow for unemployment of capital. This was done by assuming that unemployment of capital varies in direct proportion to unemployment of labor. Thus, if the unemployment rate in a certain year was 6 percent (i.e., 94 percent of the labor force was employed), the employed capital stock in that year was taken to be 94 percent of the total capital stock.

Thus, y_n is 0.027, and k_n is 0.048. In addition, the share of income going to capital (the value of α) in 1922 was 0.339. Using these figures, we have

$$r = y_n - \alpha k_n$$
$$r = 0.027 - (0.339)(0.048)$$
$$r = 0.011$$

Thus, Solow estimated that the rate of technological change in 1922 was 1.1 percent.

On the basis of calculations of this kind for each year, Solow estimated that the rate of technological change averaged about 1.5 percent per year over the period 1909–49. His most important finding was that technological change accounted for about seven eighths of the total increase in output per man-hour during this period, while capital formation accounted for only about one eighth. In other words, on the basis of this model, it would appear that technological change has been about seven times as important as investment as a source of economic growth in the United States. This would seem to suggest that measures to stimulate investment would not be very effective in raising the growth rate. Indeed, if an effort is to be made to increase the rate of growth, Solow's analysis indicates that it would be more efficient to undertake measures to accelerate technological change, since this appears to be much more important than capital formation as a stimulant to growth. It should, of course, be recognized that in the context of Solow's model "technological change" has a very broad meaning—indeed, it takes in everything that affects output per man-hour except investment in business plant and equipment. Thus, governmental measures aimed at stimulating "technological change" might include not only devices for encouraging research and invention but also expenditures on education and public health designed to improve labor skills and efficiency, expenditures on such things as highways which would indirectly raise the efficiency of private industry, and so on.

The model used by Solow in this study involves three crucial assumptions: (1) that the economy operates under conditions of constant returns to scale; (2) that factors of production receive real rates of remuneration equal to their marginal physical products, which would, strictly speaking, be the case only under conditions of pure competition; and (3) that technological change is independent of capital formation in the sense that technological improvements do not need to be embodied in new capital equipment in order to produce their effects. The first assumption is probably fairly innocuous, since there is some evidence that the American economy has, in fact, experienced constant returns to scale, at least approximately. The second assumption is much more questionable, because the economy clearly contains many monopolistic elements and market imperfections which presumably add up to a substantial departure from conditions of pure competition. The third assumption is very questionable, since it seems quite apparent from everyday observation that many—perhaps most—improvements in productive technique need to be embodied in new forms of capital equipment in order to become effective. Moreover, as we shall see below, when this assumption is changed, the conclusions reached in Solow's 1957 study are radically altered.

Embodied Technological Change

The study described above involved the extreme assumption that all technological change is disembodied—i.e., that none of it needs to be embodied in capital to become effective. In another study, published in

TABLE 19–1

Illustrative Calculation of Capital Stock Adjusted for Technical Progress

	(1) *Unadjusted Capital Stock* (*K*)	(2) *Capital stock Adjusted for 5% Embodied Technical Progress* (*J*)
Capital stock, beginning of year 1	1,000.0	1,000.0
Less: Capital consumption, year 1	100.0	100.0
Plus: Gross investment, year 1	110.0	110.0
Equals: Capital stock, beginning of year 2	1,010.0	1,010.0
Less: Capital consumption, year 2	105.0	105.0
Plus: (1) Gross investment, year 2	120.0	...
(2) Adjusted gross investment, year 2 (1.05×120)		126.0
Equals: (1) Capital stock (*K*), beginning year 2	1,025.0	...
(2) Adjusted capital stock (*J*), beginning year 3	...	1,031.0
Less: Capital consumption, year 3	108.0	108.0
Plus: (1) Gross investment, year 3	120.0	...
(2) Adjusted gross investment, year 3 ($1.05^2 \times 120$)	...	132.3
Equals: (1) Capital stock (*K*), beginning of year 4	1,037.0	...
(2) Adjusted capital stock (*J*), beginning of year 4	...	1,055.3

1962, Solow experimented with the alternative extreme assumption that all technological advance needs to be embodied in newly produced capital goods before it can have any effect on output.[20] Solow starts by assuming that capital goods produced in any year are *g* percent more productive than capital goods produced the year before. On the basis of this assumption, he constructs adjusted estimates of the capital stock corrected for this steadily increasing efficiency. His procedure is illustrated by a numerical example in Table 19–1, in which it is assumed that the capital stock at the beginning of the initial year is 1,000 and that *g*, the rate of increase in the productivity of new capital, is 5 percent per year.

[20] R. M. Solow, "Technical Progress, Capital Formation, and Economic Growth," *American Economic Review*, Vol. LII, May 1962, pp. 76–78.

The annual amounts of capital consumption allowances and gross investment are taken to be given in this example, and the unadjusted capital stock, K (column 1), is calculated for each year by subtracting capital consumption allowances from the stock at the beginning of the previous year and adding gross investment. The capital stock adjusted for technical change, J (column 2), is calculated in a similar way except that gross investment is multiplied by a factor—1.05 for year two, 1.05^2 for year three, etc.—to correct for the assumed 5 percent per year increase in the efficiency of new capital goods resulting from technological improvement.

Note particularly that the improvement factor is applied to *gross* investment, which includes investment intended to replace depreciated capital equipment as well as new investment which increases the stock of capital. When depreciated equipment is replaced, it can be assumed that the newest model is installed and therefore that replacement investment as well as new investment is a vehicle by which new technology is introduced into the economy.

Solow uses a Cobb-Douglas production function with constant returns to scale. The function takes the form

$$Y = AJ^{\alpha}N^{1-\alpha} \tag{10}$$

It differs from the production functions discussed earlier in two respects: (1) there is no allowance for disembodied technical change (i.e., the term e^{rt} is absent), and (2) the capital stock adjusted for embodied technical change (J) is substituted for the unadjusted capital stock (K). Y is defined as potential output—that is, the output that would be produced with the economy operating at 4 percent unemployment—while N is man-hours worked at full employment.[21] The analysis is limited to the private business economy; that is, the GNP data exclude the product originating in government, households, the rest of the world, and the services of houses; and the investment series excludes investment by nonprofit institutions.

Several different series for adjusted capital stock (J) are constructed, using different assumed values of g, the rate of embodied technical change. Typical are the results obtained when g takes on the value of 3 percent, that is, when it is assumed that the efficiency of new capital goods increases at 3 percent per year. Fitted by regression techniques to data for the period 1929–61, the equation yields a value of 0.51 for α, so that the growth equation becomes

[21] Since Solow fits his equation by regression techniques to data for the period 1929–61 and since the economy operated at widely varying unemployment rates during this period, he introduced an adjustment factor which makes the relation between actual and potential output depend on the unemployment rate. This adjustment factor drops out when the unemployment rate is 4 percent, leaving Equation 10 applicable as an explanation of potential output (defined as the output that the economy can produce when it is operating at 4 percent unemployment). As pointed out in Chapter 17, footnote 4, the adjustment factor enabled Solow to estimate the gap between actual and potential output as a function of the unemployment rate, and his results in this regard turned out to be remarkably similar to those of Okun, discussed earlier in that chapter.

$$y = 0.51j + 0.49n \tag{11}$$

where y is the percentage rate of growth of potential output, j is the percentage rate of growth of the adjusted capital stock, and n is the percentage rate of growth of full employment man-hours.

For purposes of contrast, Solow also estimated an equation for the same period of the type

$$Y = Ae^{rt}K^{\alpha}N^{1-\alpha}$$

in which technical change is entirely disembodied and the unadjusted capital stock is employed. This yielded a value of 2.5 percent for r and 0.11 for α. The corresponding growth equation is

$$y = 0.025 + 0.11k + 0.89n \tag{12}$$

The estimated rate of disembodied technical change of 2.5 percent derived in this study covering the period 1929–61 contrasts quite sharply with the estimate of about 1.5 percent given in Solow's earlier study covering the period 1909–49, which was summarized above.

The implications for economic growth of the two equations (11 and 12) are strikingly different. To illustrate the implications of disembodied technical change as reflected in Equation 12, let us assume that full employment man-hours grow at 1 percent per year ($n = 0.01$). Suppose that initially the capital-output ratio (K/Y) is 1.75 and that 2.5 percent of the capital stock is retired each year. On these assumptions, the investment requirements for various rates of growth are as shown in the following table:[22]

		Required Investment as Percent of Y		
(1) Rate of Growth of Potential Output (y)	*(2) Rate of Growth of Capital (k)*	*(3) Replacement Investment*	*(4) Net Investment*	*(5) Gross Investment*
3.40%	...	4.4%	...	4.4%
3.50	1.00%	4.4	1.0%	5.4
3.81	3.81	4.4	6.7	11.1
4.00	5.60	4.4	9.7	14.1
4.50	10.10	4.4	17.7	22.1

[22] The calculations in this table were made as follows:

Column 2: These figures are obtained by solving Equation 12 for k when $n = 0.01$ and y takes on the value shown in column 1.

Column 3: Depreciation (D) is 2.5 percent of capital (K); that is $D = 0.025K$. Since it is given that $K = 1.75Y$, this implies that $D = (0.025)(1.75)Y = 0.044Y$. Thus, 4.4 percent of income must be invested to make up for depreciation and hold the capital stock constant.

Column 4: If s is the saving (investment) rate, the rate of growth of capital is sY/K. That is, $sY/K = k$, or $s = kK/Y$. Since it is given that $K/Y = 1.75$, we have $s = 1.75k$, with k taking on the value shown in column 2.

According to these calculations, output will grow at 3.4 percent as a result of disembodied technical change and growth of the labor force with no growth at all in the stock of capital. Gross investment will need to be only 4.4 percent of output, enough to make up for depreciation. To raise the rate of growth to 4 percent by increasing investment will require that the fraction of output used for gross investment be increased to 14.1 percent, while a growth rate of 4.5 percent will necessitate that 22.1 percent of output be allocated to gross investment. It may be noted that the equilibrium rate of growth—the rate that will prevail if capital and income are growing at the same rate—is 3.81 percent. To achieve this rate immediately with an initial capital-output ratio of 1.75 will require that 11.1 percent of output be allocated to investment, of which 4.4 percent will be required for replacement and 6.7 percent will be net investment.[23] At a saving rate s, lower than 6.7 percent, the capital-output ratio will gradually decline; while at saving rates higher than this, the ratio will rise. In accordance with the analysis presented earlier in this chapter, if the saving rate remains constant, the rate of growth will gradually approach its equilibrium value of 3.81 percent, no matter what the saving rate is. In order to cause output to rise at a rate faster than 3.81 percent in the long run, the saving rate will have to rise gradually, that is, a continually increasing fraction of output will need to be allocated to investment.

Estimation of the capital requirements for growth when technology is embodied in capital (Equation 11) is more difficult because the adjusted capital stock grows not only as a result of net investment but also as a consequence of technical progress. However, Solow estimates the investment requirements for alternative rates of growth with embodied technical change, under conditions assumed above, as follows:

Rate of Growth (y) of Potential Output	*Required Gross Investment as % of Y*
3.5%	9.8%
4.0	10.9
4.5	12.0

Comparing the results obtained by Solow on the alternative assumptions that technology is wholly disembodied and wholly embodied, the implications of the latter model seem more reasonable. That the economy could achieve a 3.4 percent growth of potential output when the labor force grows at 1 percent per year with no net investment, as Solow's results with disembodied technical change suggest, seems wholly implausible. Moreover, it seems very unlikely that the growth rate is as little affected by additional investment as this model suggests: to increase the rate of growth from 3.5 to

[23] Using Equation 7 above, the student can verify that the *equilibrium* capital-output ratio is 1.75, the same as the assumed initial capital-output ratio, when the saving rate (s) is 6.7 percent.

4 percent would require that the fraction of output devoted to investment be nearly tripled (from 5.4 percent to 14.1 percent). The model that assumes embodied technology indicates that investment of 9.8 percent of output would be needed to achieve a growth rate of 3.5 percent and that the investment ratio would need to rise by about 11 percent (of itself), to 10.9 percent, in order to raise the rate of growth to 4 percent. If anything, however, this model may make the growth rate appear to be more responsive to changes in the investment ratio than is really the case. On purely statistical grounds, it is almost impossible to select one of Solow's equations as being superior to the other. His equations, whether they assume disembodied technical change or embodied technical change with various values of g (the rate of increase in the efficiency of new capital) all explain the facts of U.S. growth from 1929 to 1961 exceedingly well.[24]

Both Embodied and Disembodied Technical Change Present

It seems reasonable to suppose that technology is partly embodied and partly disembodied rather than entirely the one or entirely the other. This possibility has been explored by Lester C. Thurow and Lester D. Taylor in a paper published in 1966.[25] In general, the approach taken by Thurow and Taylor is rather similar to that employed by Solow in his 1962 paper. However, they do introduce one important new wrinkle: they allow for the embodiment of technical progress in labor as well as in capital. The idea here is that the quality of the labor force increases from year to year as a result of increased education, improved health, and enhanced technical skills. In addition, they allow for disembodied technical change.

Thurow and Taylor tried out various alternative formulations of their model and selected as their "preferred" equation one which, expressed in terms of rates of growth, and fitted to data for the private economy for the period 1929–65, is as follows:

$$y = 0.009 + 23j + 0.77m$$

where j is the rate of growth of the capital stock adjusted for improvement at the rate of 4 percent per year, and m is the rate of growth of full employment man-hours adjusted for improvement at the rate of 1 percent per year. As one would expect, the properties of this equation are intermediate between those of the two equations (11 and 12) above taken from Solow's 1962 paper. Disembodied technical progress occurs at a rate of 0.9 percent per year as compared with zero in Equation 11 and 2.5 percent in

[24] The values of R^2 (the coefficient of determination) range between 0.96 and 0.99 for all of Solow's equations.

[25] L. C. Thurow and L. D. Taylor, "The Interaction Between the Actual and Potential Rates of Growth," *Review of Economics and Statistics*, Vol. XLVIII, November 1966, pp. 351–60.

Equation 12. On the other hand, the elasticity of output with respect to capital is 0.23, about half the value of 0.51 in Equation 11 and about twice the value of 0.11 in Equation 12. Thus, the investment requirements for moderate rates of growth are higher than in the case in which all technology is disembodied (Equation 12) but lower than in the case in which all technology is embodied (Equation 11). On the other hand, the responsiveness of the rate of growth to changes in the rate of capital accumulation is greater than in the case in which all technology is disembodied but less than in the case in which all is embodied.

While the "preferred" equation of Thurow and Taylor seems reasonable, this equation explains the facts of U.S. growth from 1929 to 1965 only slightly better than a number of other variants they tried. These alternatives ranged from equations in which it was assumed that all technology was disembodied to equations in which it was assumed that all technology was embodied with alternative assumed rates of improvement for capital and labor. Indeed, some 36 alternative equations were tried, and the explanatory power of each of them is so great that it is not possible to single out any one as clearly superior to the rest.

While the research we have just reviewed is interesting and ingenious and a study of it brings out some of the difficulties in isolating the factors responsible for growth, the results are inconclusive and somewhat discouraging. Studies of this kind have thus far proved incapable of telling us whether technology is primarily embodied or primarily disembodied—although the former possibility seems the more intuitively reasonable—or of giving us a clear idea how potent an increase in investment is likely to be in accelerating growth.

AN ALTERNATIVE APPROACH

The most comprehensive effort to estimate the sources of past economic growth in the United States has been made by Edward F. Denison.[26] Denison does not use a production function but instead adopts a somewhat different approach to the problem. He starts with time series for total output (real GNP), labor input, capital input, and the shares of income received by labor, capital, and land for the period 1909–58. The series for output and the various inputs are converted to index numbers, with 1929 as the base year. His basic method is illustrated by the following table which shows the entries for 1929 and the calculated values for 1955. The entries in columns 1, 2, and 3 are derived by simply converting the basic data for output, labor input, and capital input to index numbers based on 1929 by dividing through by the 1929 values. Denison treats land as a third factor of

[26] E. F. Denison, *The Sources of Economic Growth in the United States and the Alternatives Before Us* (New York: Committee for Economic Development, 1962).

	Index Numbers					
Year	(1) *Total Output (Real GNP)*	(2) *Labor Input*	(3) *Capital Input*	(4) *Land Input*	(5) *Total Input*	(6) *Output per Unit of Input* [(1) ÷ (5)]
1929	100	100.0	100.0	100	100.0	100.0
1955	216	176.2	153.9	100	166.7	129.6

production, the quantity of which he assumes is constant throughout the period; thus, the entries in column 4 are simply 100 for all years. The index of total input (column 5) is a weighted average of the indexes of labor input, capital input, and land input (columns 2, 3, and 4) with the relative income shares of labor, capital, and land for the period 1909–58 used as weights. According to Denison's calculations, these shares are as follows:

	Percent of National Income Received 1909–58
Labor	71.4
Capital	22.8
Land	5.8
Total	100.0

Thus, the entry for 1955 in column 5 is calculated as follows:

$$0.714 \times 176.2 + 0.228 \times 153.9 + 0.058 \times 100 = 166.7$$

The idea here is that each factor is weighted in accordance with its economic importance as measured by the share of national income it receives.[27] The last step in the calculation is to divide the index number of total output (column 1) by the index number of total input (column 5) to obtain an index number for output per unit of input (column 6). Thus, the index number of 129.6 for output per unit of input in 1955 is obtained by dividing the index number for total output (216) by the index number for total input (166.7).

The data presented in the table of index numbers can be used to disentangle the contributions of factor inputs and of output per unit of input to total economic growth. The increase in the index number for total output from 100 to 216 during the 26-year period 1929 to 1955 works out to an average increase of 3 percent per year, compounded annually. Similarly the increase in total input (column 5), calculated in the same way, comes out to 2 percent per year, while the increase in output per unit of input (column 6) amounts to 1 percent per year. That is, of the total growth of 3 percent per

[27] It may be noted that in his final calculations Denison does not use the income shares for the entire period 1909–58 as weights; rather he changes the weights every five years in order to keep them more closely up to date. This refinement does not, however, alter the essential nature of his calculations.

TABLE 19–2

Allocation of Growth Rate of Real GNP among the Sources of Growth, 1929–57

	Percentage Points in Growth Rate	*Percent of Growth Rate*
Real GNP—total growth	2.93	100
Increase in total inputs	2.00	68
Labor, adjusted for quality change	1.57	54
Employment	1.00	34
Change in hours of work	−0.20	−7
Education	0.67	23
Other factors*	0.10	4
Capital	0.43	15
Increase in output per unit of input	0.93	32
Economies of scale	0.34	12
Advance of knowledge	0.59	20

NOTE: Items may not add to totals due to rounding.

* Increased experience and better utilization of women workers and changes in age-sex composition of the labor force.

SOURCE: E. F. Denison, *The Sources of Economic Growth in the United States and the Alternatives Before Us* (New York: Committee for Economic Development, 1962), Table 32, p. 266.

year in output, two thirds, or 2 percent, is accounted for by the increase in total factor input, while the remaining one third, or 1 percent, is accounted for by increased output per unit of input.[28]

This procedure thus enables Denison to isolate the contributions to economic growth resulting from (*a*) increases in labor input, (*b*) increases in capital input, and (*c*) increases in output per unit of input or factor productivity. Since he is interested in growth of *potential* output, he selects periods in which economic conditions—particularly the unemployment rate—were at least approximately the same at the beginning as they were at the end. Although he utilizes several such periods, he places greatest emphasis on the period 1929–57, and we will examine his conclusions for that period only.

Denison's results for the period 1929–57 are summarized in Table 19–2.[29] The index number for real GNP increased from 100 in 1929 to

[28] The contribution of the increase in total input can be further subdivided into the contribution attributable to labor input and that attributable to capital input. Thus, labor input (column 2) grew at 2.2 percent per year while capital input (column 3) grew at 1.7 percent per year. Weighting these growth rates by the importance of capital and labor in production as measured by proportions of national income received, we obtain 1.6 percent (2.2 percent × 0.714) for labor and 0.4 percent (1.7 percent × 0.228) for capital.

[29] In the interest of simplicity, some of the details of Denison's findings have been omitted. He breaks down the contribution of capital formation according to subtypes of capital such as nonfarm residental structures, other structures and equipment, and so on; this breakdown has been omitted from Table 19–2. In addition, some minor factors affecting growth which net out to zero (because some tended to increase growth and others to reduce it) have been left out.

224.8 in 1957; this works out to an average increase of 2.93 percent per year during the 28-year period. The calculations summarized in Table 19–2 indicate that 2 percentage points—or 68 percent—of this growth is accounted for by the increase in total inputs, while the remaining 0.93 percentage point—or 32 percent of the total—is attributable to the increase in output per unit of input. Of the 2 percentage points accounted for by the increase in total inputs, 1.57 percentage points of growth is attributable to increased labor input, while the remaining 0.43 percentage point is attributable to growth of the capital stock.

One of the most interesting features of Denison's study is the fact that he makes an attempt to estimate the contribution of education to economic growth and incorporates this estimate in his calculation of labor input; that is, his estimate of labor input is adjusted for quality change attributable to improved education. As a basis for adjusting the labor input for quality improvement and calculating the contribution of education to economic growth, Denison relies primarily on data drawn from the 1950 census which enables him to estimate income differentials attributable to education in 1949. Assuming that these income differentials are a satisfactory measure of differential contributions to production of persons with different levels of education and projecting forward and backward from 1949 on the basis of estimates of the distribution of the labor force by educational level in other years, Denison is able to complete his adjustments of the labor force for quality and his estimates of the contribution of education to economic growth. According to his estimates, the contribution has been a major one; Table 19–2 indicates that education accounted for 0.67 percentage point of growth per year—or 23 percent of the total—during the period 1929–57.

The increase in output per unit of input, which, according to Denison, accounted for 0.93 percentage point of growth per year from 1929–57, is essentially a residual element in his basic calculations, that is, it is simply that portion of the growth rate that is not accounted for by increased factor inputs. However, Denison makes an effort to break this residual down into meaningful components. A number of these components which are rather trivial and which, on balance, cancel out (because some increase the growth rate and others reduce it) have been omitted from Table 19–2 in the interest of simplicity. The one factor that is important is economies of scale. Denison concludes that the U.S. economy is subject to increasing returns to scale in both national and local markets—that is, that a 1 percent increase in inputs of all factors of production will cause total output to increase by more than 1 percent. For the period 1929–57, he estimates that such economies of scale were responsible for 0.34 percentage point of growth per year or 11 percent of the total growth rate.

When the 0.34 percentage point of growth attributed to economies of scale is subtracted from the 0.93 percentage point for the increase in output per unit of input, a remainder of 0.59 percentage point is arrived at. This portion of growth, which amounts to 20 percent of the total growth rate, is attributed by Denison to the advance of knowledge. This is the final residual in the calculations.

Advance of knowledge, in this context, means improvements in technology or in managerial practices which permit existing goods and services to be produced more efficiently—that is, with smaller inputs of labor and capital per unit of output. It will be recalled that real GNP as a measure of growth suffers from the defect that it does not fully reflect improvements in the quality of products and is almost completely unaffected by the introduction of wholly new products. To the extent that new inventions and technological changes result in improved quality of products or in the introduction of new products, they have little effect on growth as measured by real GNP. Since a large portion of technological advance does undoubtedly take the form of quality improvements and the introduction of new products, the estimate of 0.59 percentage point substantially underestimates the contribution of the advance of knowledge to true economic growth. Of course, the total growth rate of 2.93 percent for real GNP correspondingly underestimates the true growth of the economy.

It should be noted that Denison attributes no part of recorded economic growth to investment by federal, state, and local governments in such public-sector activities as highways, urban development, public housing, and so on. The reason for this omission is that the national income accounts do not recognize these activities of governments as constituting capital formation. No estimates are made of the capital stock in the public sector, and no allowance for the value of the services of such capital is included in the national income or product. For this reason, such investment does not add to the productive capacity of the economy as measured by potential GNP. Investments in the public sector affect the measured growth of GNP only indirectly, to the extent that they lead to increased efficiency in the private sector—for example, an improved highway system, by providing better transportation facilities, may make private business more productive. The major contribution of the highways in providing consumers with greater opportunities for travel and improved access to recreational facilities is entirely omitted from the GNP.

The main factors responsible for growth during the period 1929–57, according to Denison, may be summarized as follows:

	Percentage Points of Growth	*Percent of Total Growth Rate*
Labor input, unadjusted	0.90	31
Capital input	0.43	15
Education	0.67	23
Economies of scale	0.34	12
Advance of knowledge	0.59	20
Total Growth Rate	2.93	100

Perhaps the most surprising finding in Denison's study is the relatively small contribution of capital formation, which is indicated as accounting for only 15 percent of the growth rate. Both education and the advance of knowledge are estimated to have made substantially larger contributions to growth than this.

Denison's study contains many valuable insights into the process of economic growth, and his quantitative breakdown of growth by sources is useful as at least a crude approximation. However, the crudity of his estimates should be emphasized, lest they be taken too seriously. At a number of points—for example, with respect to the contribution of education to growth—the data are scarcely adequate to support satisfactory estimation. At certain critical points in the study, Denison was forced to resort to pure guesswork. It will suffice to mention two instances of this. First, in his treatment of economies of scale, after some discussion of the nature of such economies and the opinion of other economists regarding them, Denison winds up by simply assuming quite arbitrarily that economies of scale accounted for 12 percent of total economic growth during the period 1929–57. Second, in making his estimates of the contribution of education to economic growth, Denison was faced with the obvious problem that the differentials in income that are *associated with* differences in education are not entirely *caused by* these differences. To some extent, persons who possess the inherent advantages of unusual ability, energy, and motivation are likely to seek and obtain more education than other people. Denison's solution to this problem is to assume, again quite arbitrarily, that 60 percent of the differentials in income associated with differences in education are caused by these differences.[30]

CONCLUSION

We may conclude that while the work of Solow, Denison, and others provides considerable insight into the process of economic growth, we are still far from having firm knowledge of the relative importance of various factors, such as capital formation, education, and technical change, in generating growth. Indeed, we may say that the problem of growth has been delineated but not solved.

[30] For an excellent summary and critique of the Denison study, see Moses Abramovitz, "Economic Growth in the United States: A Review Article," *American Economic Review*, Vol. LII, September 1962, pp. 762–82.

Chapter 20 GROWTH POLICY

As indicated in Chapter 17, when the economy is operating below full employment, it may be possible to achieve rapid growth in the short run merely by adopting policies to expand aggregate demand, thereby bringing idle productive capacity into use. Moreover, the achievement of full employment is itself likely to have favorable effects on some of the factors determining the growth of capacity. In this connection, two factors are likely to be especially important.

1. The improvement in rates of utilization of existing productive facilities associated with the achievement of full employment is very likely to provide a stimulus to private investment in new plant and equipment, thereby leading to faster growth of the stock of capital. And, since the installation of equipment is often the vehicle by which new technology is introduced into productive processes, the allocation of more resources to investment should also lead to more rapid introduction of new techniques.

2. Unemployment and underutilization of productive facilities increase incentives to seek job security and the protection of limited markets through the adoption of restrictive practices which cut down the effective productive capacity of the economy and reduce its rate of growth. Labor unions seek to protect their members by engaging in "featherbedding" and resisting the introduction of improved production techniques which may eliminate jobs. Businessmen are motivated to adopt pricing policies designed to protect their existing markets and to avoid risky new ventures which might seem attractive in a period of strong and expanding markets.

Undoubtedly, the particular kinds of measures that are used to expand aggregate demand may have some effect on the fraction of national resources devoted to investment during the course of an expansion to full employment and therefore on the growth of productive capacity. For example, if the expansion is powered by an easy monetary policy which reduces interest rates, expanded investment may play a more important role in the expansion than will be the case if the recovery is caused by a reduction in personal income taxes to stimulate consumption or by an increase in government expenditures. But when the economy has been operating below full employment for some time, the most important factor in strengthening the inducement to invest in new plant and equipment is an expansion of aggregate demand which will increase utilization rates of existing facilities. That is, in such circumstances, the generation of additional aggregate demand is much more important for growth than the precise nature of the measures adopted to achieve this objective, even though these specific measures may have a second-order influence on the size of the investment component of the expansion.

INCREASING THE RATE OF GROWTH UNDER FULL EMPLOYMENT CONDITIONS

Suppose it is decided that growth is not rapid enough at a time when the economy is already operating at full employment, and that government action should be taken to accelerate growth. Under these circumstances, a mere expansion of aggregate demand will not be capable of increasing the rate of growth without generating excessive inflation. Acceleration of growth under full employment conditions requires that measures be taken to expand productive capacity. As capacity expands, aggregate demand will, of course, have to grow in pace if full employment is to be maintained. But the first step in the process must be an increase in capacity.

One way to expand productive capacity is to increase the efficiency with which the economy operates. An especially promising way to increase efficiency is through the adoption of policies to improve the working of the labor market. Increased appropriations for the United States Employment Service, to enable that agency to expand its activities in disseminating job information for the benefit of workers seeking employment; the use of federal subsidies or perhaps more generous tax credits for moving expenses under the individual income tax, to reduce the cost to workers of moving from one locality to another to accept employment; and expanded federal programs for the training and retraining of the unemployed would all be helpful in increasing labor mobility and permitting a more rapid adaptation of supply and demand in particular labor markets with less upward pressure on wages and prices. Combined with measures to strengthen competition and increase the downward flexibility of industrial prices and to prevent inflationary excesses in collective bargaining, measures of this kind might serve to shift the modified Phillips curve (Figure 17–1) to the left, thereby permitting any given rate of unemployment to be achieved with less inflationary pressure than formerly. This might allow full employment to be redefined to be, let us say, 3 percent rather than 4 percent unemployment, thereby permitting the capacity growth curve (Figure 17–2) to be redrawn at a higher level. Thus, while not necessarily increasing the *percentage rate of growth* of productive capacity, policies to increase efficiency of labor markets would, if effective, produce a "once-and-for-all" increment to capacity. Since they would enable the economy to produce larger quantities of goods and services not only currently but in the future, these measures should be classified as growth policies. And, since they would reduce the economic distress and disillusionment created by unemployment, they would be valuable measures from the standpoint of social policy as well.

Except for policies such as these, which increase output by enabling the economy to function more efficiently, virtually all measures for accelerating growth under conditions of full employment involve the reallocation of resources from uses that do not contribute to growth to uses that do so contribute. That is to say, they require that current consumption be reduced and that the resources released by this reduction be channeled into some form of investment. Thus, the policies adopted for this purpose must be such as to keep aggregate demand approximately unchanged at the full employment level while shifting the composition of demand away from consumption and toward investment. Or, to put it another way, total saving—including government saving through the budget—would have to be increased, and this increase in saving would have to be matched by an increase in investment.

To give a general idea of the kinds of policy actions that might be taken to stimulate growth, the government might increase personal income tax rates; this would increase tax collections at the current income level, thereby reducing disposable income and depressing personal consumption expenditures. If tax revenues were increased by $5 billion and the marginal propensity to consume was 90 percent, this would reduce consumption spending by $4.5 billion. It would increase total saving in the economy by a like amount, since it would shift the federal budget in the direction of a surplus by $5 billion while reducing personal saving by $0.5 billion. To offset the effect of the tax increase on aggregate demand, the government would have to take action to increase growth-generating expenditures by $5 billion. Measures that might be taken to accomplish this purpose would include:

1. An easing of monetary policy by the Federal Reserve to lower interest rates for the purpose of inducing businessmen to increase investment in new plant and equipment, or the adoption of tax incentive devices to accomplish the same purpose.
2. Increased federal spending—either directly or through grants to state and local government units—on education and training or other forms of investment in human capital such as health programs.
3. Increased federal spending in support of the search for knowledge through research and development, or the introduction of new incentives through subsidies or tax credits to stimulate additional private expenditures for this purpose.
4. Additional federal spending—either directly or through grants to state and local governments—on investment in public-sector activities that

contribute to growth, such as highway construction, urban development, and development of natural resources.

As these measures to increase growth came to fruition in the form of increased productive capacity, measures would, of course, have to be taken to insure a sufficient growth of aggregate demand to utilize the added capacity in order to maintain full employment.[1]

It is almost certain that the optimal way to increase economic growth is by the use of some combination of the measures listed above: to allocate some additional resources to private investment, some to the improvement of human resources, some to the expansion of research and development activities, and some to increased public investment in physical facilities.[2] The general principle that should underlie selection of the optimal mix of policies for achieving a given growth objective is to carry each of the various growth-generating activities to the point where the marginal social productivities of all of them are equated. Then, having decided the optimal combination of these activities for each given increment to the growth rate, the total amount of resources withdrawn from current consumption for use in promoting growth-oriented activities should be decided on the basis of our willingness as a nation—as reflected in the decisions of our policy makers chosen through democratic political processes—to give up current consumption in exchange for future consumption.

Unfortunately, while it is a relatively simple matter to state, at least crudely, the principles that should underlie the selection of an optimal growth policy, it is, as a practical matter, impossible to make such a rational calculation in the present state of knowledge. As has already been brought out in Chapter 19, opinions differ substantially concerning the relative magnitudes of the contributions to economic growth that have been made by private investment, technological change, education, and so on; and it is

---

[1] For a formal discussion of the use of policy instruments to maintain a balance between the growth of capacity and the growth of demand in the context of a Domar-type model, see W. L. Smith, "Monetary-Fiscal Policy and Economic Growth," *Quarterly Journal of Economics*, Vol. LXXI, February 1957, pp. 36–55. See also W. H. L. Anderson and John Cornwall, "Problems of Growth Policy," *Review of Economics and Statistics*, Vol. XLIII, May 1961, pp. 163–74; and R. R. Nelson, "Full Employment Policy and Economic Growth," *American Economic Review*, Vol. LVI, December 1966, pp. 1178–92.

[2] As was pointed out in the introduction to this part of the book, to the extent that research and development activities result in new consumer products or improvements in the quality of existing products, their effects show up only to a very minor extent in the measured growth of GNP. Likewise, the effects of government investment in physical facilities, such as highways, urban development, and so on, are not reflected in the growth of GNP because it is virtually impossible to estimate the returns to such investments. Nevertheless, there can be no doubt that research that leads to new and improved products and investments that occur in the public sector make important, even though unmeasurable, contributions to the improvement of living standards. Our inability to measure their contributions should not lead us to neglect these obviously important activities as part of a program to accelerate growth.

possible to marshal the evidence in such a way as to support a fairly wide range of estimates with respect to these contributions.[3] Of course, further empirical work on the sources of economic growth may in time enable us to make better judgments concerning the contributions of different kinds of growth-promoting activities. For the present, however, about the best that can be said is that a combination of measures aimed at all of the main sources of economic growth simultaneously is probably better than single-minded concentration on one source, such as private investment. But, with respect to the relative emphasis to be placed on different kinds of activities, the judgments of qualified students of growth differ, and no clear choice seems possible.

SHOULD THE GOVERNMENT HAVE A CONSCIOUS GROWTH POLICY?

We have discussed various ways in which the federal government might be able to adjust its policies in order to change the rate of growth while still keeping the economy operating at full employment. Before concluding our discussion, we should consider the question: Is it, in fact, proper for the federal government to attempt to influence the rate of growth by its policies? This is a question concerning which there is by no means unanimous agreement among economists, although the answers given to it often involve value judgments and ideology to a greater extent than they do economic analysis.[4]

The issue of economic growth is basically one that involves the relation

---

[3] As a part of his study of the sources of economic growth, Edward F. Denison (*The Sources of Economic Growth in the United States and the Alternatives Before Us*, [New York: Committee for Economic Development, 1962]) prepared a "menu" of some 31 measures that might be taken to raise the growth rate, together with (for most of the measures) quantitative estimates of the contributions they might be capable of making. Drawing on this "menu," he also put together a collection of 13 measures which he estimated would be capable in combinations of raising the growth rate by 1 percentage point per year for the period 1960–80 (see his paper, "How to Raise the High-Employment Growth Rate by One Percentage Point," *American Economic Review*, Vol. LII, May 1962, pp. 67–75). Many of the items on Denison's menu are essentially "one-shot" changes; that is, once done they could not be repeated. For example, he estimates that racial discrimination costs us 0.8 percent of the national income. If discrimination could be abolished in the course of the 20-year period, it would raise national income by 0.8 percent or 0.04 percent per year (0.8 ÷ 20). Moreover, many of the items on Denison's list—such as abolition of racial discrimination—are not changes in policy, but rather are changes in the structure and operation of our economy and society which would require complex, forceful, and many-faceted policy actions to bring about. Finally, while Denison exhibited great resourcefulness and ingenuity in making his estimates, his methods were necessarily crude, and the data available to him were in some instances not satisfactory. As a consequence, his results would certainly not be accepted as reliable by many students of economic growth.

[4] For an excellent discussion of this issue, see James Tobin, "Economic Growth as an Objective of Government Policy," and the comments on Tobin's paper by H. G. Johnson and Herbert Stein, *American Economic Review*, Vol. LIV, May 1964, pp. 1–27.

between the present and the future—or, at least roughly, the relations between the present generation and future generations. As we have seen, under conditions of full employment, most of the measures that might be taken by the government to increase growth require that current consumption be reduced and that the resources thereby released be channeled into some form of investment activity. Thus, the basic issue of growth policy is: Shall measures be taken by the government to reduce the consumption of the present generation in order to permit some future generation or generations to enjoy higher consumption?[5]

Private individuals and business enterprises make provision for the future through personal and business saving. There is general agreement that the federal government should take action to ensure that there is sufficient demand over and above private consumption to offset the private saving that is forthcoming at full employment. If this is not done, full employment will not be maintained, resources will go to waste, and, to some degree, the saving proclivities of the private sector will be aborted as a result of a decline in income.

One view concerning the growth question is that the saving of the private sector represents the provision for the future that the citizenry of the nation is willing to make voluntarily, balancing its desires for current consumption against the higher income (and consumption) that can be obtained in the future through the process of saving and investing. The growth that results from this process constitutes the provision for future generations that citizens of the country, acting in their private capacities, are willing to make. It is argued that the decisions of the private sector in this regard should be respected by the government; thus, apart from seeing that the full employment saving of the private sector is matched by an equal amount of investment, the government has no other role to play in influencing growth. To put it simply, according to this view, the government should be neutral with respect to growth, leaving the matter to be determined by the decisions of the private sector.

In a country which relies primarily on markets and the tastes of its citizens to govern the allocation of resources, there is a certain appeal in the idea of allowing markets and the tastes of individuals to determine the provision that is made for the future. The trouble is that it is virtually impossible, in practice, for the government to be neutral with respect to growth; its activities designed to maintain economic stability and to meet the needs of its citizens in a complex industrialized society will unavoidably affect the rate of growth of the economy, whether or not this is one of the

[5] The issue could arise in the opposite form: Should the government take measures to *reduce* the rate of growth, that is, to increase the consumption of the present generation at the expense of future generations? In practice, however, the whole discussion has had to do with measures to accelerate growth.

conscious objectives of a government. One means by which, it is sometimes suggested, the government might attempt to maintain a posture of neutrality with respect to growth would be through the adoption of a capital budget. Under such a system, the government would divide its expenditures into two categories: government consumption (expenditures yielding benefits only in the current year) and government investment (expenditures yielding benefits in the future). It would aim each year to raise sufficient tax revenues to cover its consumption expenditures, together with depreciation of its existing capital and interest on its outstanding debt. All of its new investment expenditures would then be financed by borrowing. Under such a system, net private saving would set the limit on total net investment in the economy, and the government would have to compete in the capital markets with private investors to borrow the funds needed to finance its investment activities.

The major difficulty with such a capital budget arrangement is that it would severely limit the flexibility of the government in using fiscal policy for economic stabilization. If a decline in the inducement to invest—a shift to the left of the marginal efficiency of investment schedule—threatened to bring on a recession, the government could (*a*) adopt an easy monetary policy in an effort to reinvigorate private investment, and/or (*b*) increase its own investment expenditures to fill the gap. Its ability to make adjustments in the tax rates or in its consumption expenditures would, however, be extremely limited because of the dictum that government consumption must be covered by tax revenues. There is reason to doubt whether monetary policy is sufficiently effective to be able to carry the major burden of stabilization, and changes in government investment expenditures would be difficult to make with sufficient rapidity to take up the slack. Indeed, countercyclical changes in tax rates are now viewed by many economists as an especially useful stabilization instrument; and the capital budget would effectively rule out such changes.[6]

Thus, it seems inevitable, as a practical matter, that government policy will affect the rate of growth. The question therefore becomes: Should the government disregard the effects of its policies on economic growth, just as it disregards many of the other side effects of its policies, or should it have a conscious objective—or target—with respect to growth?

---

[6] If the dictum that tax revenues should cover government consumption were adhered to every year regardless of economic conditions, tax rates would have to be *increased* when economic activity declined, because with given tax rates a decline in income would cause tax revenues to fall. Thus, tax policy would have to be conducted in a cyclically perverse manner, thereby rendering the economy more unstable. This difficulty could be avoided by specifying that the tax revenues that would be collected if the economy were at full employment should be sufficient to cover government consumption. However, changing the rule in this way would still not give appreciable leeway for independent adjustments in tax rates for stabilization purposes.

It should be noted that in a democratic society, the basic political decisions are made at the ballot box, and these decisions are just as legitimately a representation of the public will as decisions recorded in the marketplace. Thus, if the public elects a President and a Congress that are on record as favoring a national growth policy, there is little basis for arguing that such a policy contravenes the will of the people as reflected in their decisions to engage in private saving. The real issue therefore is: What reasons—if any—are there for an intelligent electorate to conclude that the national interest is served by a vigorous government policy of stimulating growth?

Two reasons that are essentially political have sometimes been advanced. One is that if we are to win the uncommitted and largely underdeveloped nations over to our political and economic system, we must demonstrate to them that our economy is capable of growth at a rate as fast as, or faster than, that of the Soviet Union. This is not a very convincing argument. There are many other aspects of our national life and of our economic performance that are surely at least as important as our rate of growth in the eyes of the underdeveloped world. Moreover, there are other nations, such as Japan, which are better object lessons of the ability of a free economy to promote rapid economic development than the United States in its present advanced stage of development can possibly hope to be. The other reason sometimes given for government policies to stimulate growth is that the survival of the nation in some future international conflict may depend on the productive capacity of the economy and that it is the government's responsibility to ensure that our capacity is sufficient to meet such a test. This, too, seems a dubious argument for emphasizing growth as an objective in a period when the horrendous developments of the nuclear age have vastly reduced the significance of a nation's total economic strength as a factor affecting its ability to survive.

There is an economic argument for a government policy designed to stimulate growth on the ground that some activities that contribute to growth will not be pursued as vigorously as they should be if their implementation is left entirely to private individuals and business firms, with no action by the government. The full benefits to society of some growth-promoting activities are not capable of being appropriated by the private sponsors of such activities—that is, the social benefits of the activities exceed the benefits accruing to the private individuals or firms undertaking them. If activities of this kind are left entirely to profit-maximizing private economic units, they will be carried only to the point where their marginal cost is equal to the marginal private benefit accruing to the units undertaking them, whereas from the social point of view further resources should be allocated to them until their marginal cost is equated to their marginal social benefit. Activities that fall in this category include research—especially basic research, where the area of potential applicability is much wider than the range of markets covered by even the largest firms.[7] Another activity of this kind is apprenticeship or on-the-job training programs conducted by private firms, where there is no assurance that the employee who is being trained will continue to work for the firm providing the training. Programs to stimulate such activities as these, either through direct government participation or through subsidies or favorable tax treatment, are appropriate as a means of encouraging the allocation of sufficient additional resources to these areas to bring marginal costs to equality with marginal social benefits. Investment in areas such as highways, urban development, and general education, which have been assigned in whole or in part to the public sector, is an important determinant of growth and is, by definition, a government responsibility in the United States. Of course, the justification for these forms of investment can rest only partially on their contribution to growth. But the government's responsibility for investment decisions in these areas does provide some economic rationale for a growth policy on the part of the government.

The federal government is already involved in the support of research and manpower training programs and in investment in highways, urban development, education, and other public-sector activities that contribute to growth. Accepting the fact that the government is inextricably involved in economic growth and that there is some rational economic basis for this involvement, one may ask how the government should formulate its growth objective. Should it set a specific target for the rate of growth and attempt to adjust its policies to achieve this target? If there is to be such a target, how should it be determined?

Some efforts have been made to determine an optimal saving rate and a corresponding optimal growth path for the economy. This can be done for an economy whose growth is governed by a Cobb-Douglas production function with neutral technological change (such as Equation 4 in Chapter 19). We have

$$Y = Ae^{rt}K^{\alpha}N^{1-\alpha}$$

It was shown in Chapter 19 that when the rate of growth (n) of the labor force is given, the equilibrium growth rate for such a model is independent of the saving rate but that the equilibrium capital-output ratio is (Equation 7, Chapter 19)

$$\frac{K}{Y} = \frac{s(1-\alpha)}{n(1-\alpha)+r}$$

[7] See R. R. Nelson, "The Simple Economics of Basic Scientific Research," *Journal of Political Economy*, Vol. LXVII, June 1959, pp. 297–306; also R. R. Nelson, M. J. Peck, and E. D. Kalachek, *Technology, Economic Growth, and Public Policy* (Washington: Brookings Institution, 1967).

The higher the saving rate, the larger the equilibrium capital-output ratio and the higher the equilibrium growth path (see Figure 19–1 and the related discussion).

For an economy whose growth is governed by such a production function, it can be shown that there is an optimal saving rate (s) in the special sense that once the economy has achieved the equilibrium capital-output ratio and growth path corresponding to that rate, it is the one out of all saving rates that if held constant from generation to generation will maximize consumption for the current and all future generations. It turns out that the optimal saving rate in this sense satisfies the relation

$$s = \alpha$$

That is, the optimal saving rate is equal to the elasticity of output with respect to capital.[8] Since virtually all the empirical evidence strongly sug-

[8] This rule is derived for a more general class of production functions, of which the Cobb-Douglas function is a special case, in E. S. Phelps, "The Golden Rule of Accumulation: A Fable for Growthmen," *American Economic Review,* Vol. LI, September 1961, pp. 638–42. For the Cobb-Douglas model, the derivation is as follows: If the saving rate is s, consumption (C) is given by

$$C = (1 - s)Y$$

If the labor force is growing at a steady rate (n) and the stock of capital is optimally adjusted to income for the specified saving rate, we have

$$N = N_0 e^{nt}$$

and

$$K = \frac{s(1-\alpha)}{r + n(1-\alpha)} Y$$

Substituting these expressions in the production function, we obtain

$$Y = Ae^{rt}\left(\frac{s(1-\alpha)}{r + n(1-\alpha)}\right)^{\alpha} Y^{\alpha} N_0^{1-\alpha} e^{n(1-\alpha)t}$$

Solving explicitly for Y, we have

$$Y = Ae^{\frac{rt}{1-\alpha}}\left(\frac{s(1-\alpha)}{r + n(1-\alpha)}\right)^{\frac{\alpha}{1-\alpha}} N_0 e^{nt}$$

or

$$C = (1 - s)Ae^{\frac{rt}{1-\alpha}}\left(\frac{s(1-\alpha)}{r + n(1-\alpha)}\right)^{\frac{\alpha}{1-\alpha}} N_0 e^{nt}$$

Differentiating this expression with respect to s and setting the derivative equal to zero as a condition for a maximum, we obtain (after some simplification),

$$\frac{dC}{ds} = -\left(\frac{s(1-\alpha)}{r + n(1-\alpha)}\right)^{\frac{\alpha}{1-\alpha}} + \frac{\alpha(1-s)}{r + n(1-\alpha)}\left(\frac{s(1-\alpha)}{r + n(1-\alpha)}\right)^{\frac{\alpha}{1-\alpha}-1} = 0$$

Multiplying through by $\left(\frac{s(1-\alpha)}{r + n(1-\alpha)}\right)^{1-\frac{\alpha}{1-\alpha}}$ and simplifying, we obtain the condition

$$s = \alpha$$

gests that the saving rate in the United States is substantially less than the elasticity of output with respect to capital, this analysis might seem to suggest that it would be appropriate for the government to take action to raise s to equality with α.

Unfortunately, no such simple answer can be given to the growth policy question. The problem is that the specified saving rate is the optimal one for the current generation only on the assumption that the economy is already on the equilibrium growth path corresponding to that saving rate. Starting from the present situation, in which the capital-output ratio is substantially below the level corresponding to the optimal path, the present generation would have to engage in a burst of saving—at a rate much higher than the "optimal" one—in order to get the economy onto the optimal path. In effect, the present generation would have to make a heavy sacrifice for the benefit of future generations in order to get the scheme started, and there is no apparent reason why it should be willing to do so. Second, even if the economy were already on the optimal path, the saving rate corresponding to that path is the optimal rate only on the assumption that the present generation enters into an agreement with future generations (yet unborn!) to hold the saving rate constant over time. Obviously, any particular generation could reduce the saving rate below the optimal level and thereby increase its own consumption by violating the compact. And it is difficult to see why any generation should feel an obligation to abide by an agreement to which it was not initially a party. For these and other reasons, this concept of an optimal saving rate appears to be an interesting plaything rather than a precept about optimal growth policy that has practical applicability.

We must conclude that while public policy inevitably affects the growth rate and that it is proper for the government to recognize these effects and take them into account, any target the government sets for the growth rate must necessarily be, to a considerable extent, arbitrary and incapable of detailed defense on the basis of economic analysis.

PART V

National Income, International Trade, and the Balance of Payments

INTRODUCTION

International trade links together the economies of the trading countries, making income and employment in each country dependent on conditions not only in that country but in the rest of the world as well. Moreover, trade itself is influenced by changes in income levels in the participating countries. Up to now, we have assumed a closed economy—that is, an economy which has no trade relations with the rest of the world—and have therefore avoided the complications resulting from the interaction between income and trade. It is now time to remedy this defect in our analysis.

The plan of this part of the book is as follows. In Chapter 21, we explain the working of the international monetary system as it exists today and discuss the balance-of-payments accounts, with special emphasis on U.S. accounts. Chapter 22 examines in some detail the connection between the domestic economy and international trade, derives the impacts on income and the balance of payments of autonomous shifts in expenditures in an open economy, and discusses the policy implications of these relationships. In Chapter 23, the various alternative ways of dealing with balance-of-payments deficits and surpluses are taken up, and the conflicts that may arise between the domestic and international objectives of economic policy are analyzed. Chapter 24 draws on the material developed in the earlier chapters to analyze some of the weaknesses that exist in the present international monetary system and discusses a variety of proposals that have been advanced for improving the system.

Chapter 21 THE INTERNATIONAL MONETARY SYSTEM AND THE BALANCE OF PAYMENTS

The relationship between national income and international trade depends on the kind of international monetary arrangements that prevail. Consequently, in order to place our analysis within a reasonably realistic institutional framework, it will be necessary to begin with a description and explanation of the international monetary system as it now exists.

THE INTERNATIONAL MONETARY SYSTEM

The foundation of our present international monetary institutions was laid with the establishment of the International Monetary Fund in 1946. The membership of the IMF numbers over 100 countries. This includes nearly all countries outside the communist bloc, the only important exception being Switzerland, which, while not formally a member of the IMF, cooperates closely with it.

Exchange Parities under the IMF

The par values of the currencies of all members of the IMF are defined in terms of gold. Thus, the U.S. dollar is defined as 0.888671 grams of gold, which determines the official price of gold at $35 per troy ounce.[1] The British pound sterling is similarly defined as 2.13281 grams of gold. Since the pound contains 2.4 times as much gold as the dollar ($2.13281 \div 0.888671 = 2.4$), the par value of the pound expressed in terms of dollars is $2.40. In a similar way, exchange parities of other currencies in terms of dollars are (as of December 1969): for Germany, one deutsche mark = $0.273; for France, one franc = $0.180; for Mexico, one peso = $0.08; for Uganda, one shilling = $0.14; for Pakistan, one rupee = $0.21; and so on.

[1] One troy ounce contains 31.103481 grams; 31.103481 divided by 0.888671 equals 35.

Settlement of Private Transactions

Private international transactions in goods, services, and securities are settled by means of payment in national currencies. For example, if a French firm imports goods from Germany, payment may be made either in francs, in deutsche marks, or in the currency of a third country, such as U.S. dollars. If payment is made in francs, the French importer may remit a bank draft denominated in francs to the German exporter, who will sell the draft to his bank for deutsche marks. If payment is made in deutsche marks, the French importer may buy with francs a draft denominated in deutsche marks and remit the draft to the German exporter, who will deposit it in his bank. If payment is made in dollars, the French importer will buy dollars in the foreign exchange market and remit them to the German exporter, who will sell them for deutsche marks.[2] In the years following World War II, most countries, with the notable exception of the United States, employed exchange controls to regulate international transactions in order to conserve scarce foreign exchange for uses that were vital for domestic reconstruction. That is, exporters who received payments in foreign currencies were generally required to sell these currencies to a government exchange-control authority for domestic currency, and those wishing to obtain foreign currencies to pay for imports could get them from the exchange-control authority only if the imports were judged to serve an important national purpose. Foreigners who came into possession of a country's money were commonly restricted in their freedom to use it to buy foreign goods or to convert it into currency. However, as the needs of reconstruction were met and necessary reserves of foreign exchange and gold were built up, the industrial countries gradually relaxed their exchange controls. Indeed, since 1958, Canada, Japan, the United Kingdom, and the countries of Western Europe have had convertible currencies and relatively free foreign exchange markets in which private transactions are permitted with few restrictions. Generally speaking, transactions involving trade, shipping, insurance payments, etc., are completely free, while restrictions are often imposed on tourism and capital movements. However, most of the less developed countries of Asia, Africa, and Latin America still have extensive exchange controls which they use to conserve their limited foreign exchange for uses conducive to economic development. As a result of these restrictions as well as a general lack of international confidence in the currencies of the less developed countries,

---

[2] The arrangements for international financing and payment are quite complex in detail, but the complications need not concern us for our present purpose. For a readable exposition of the institutional arrangements, see Richard Ward, *International Finance* (Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1965); see also A. R. Holmes and F. H. Schott, *The New York Foreign Exchange Market* (Federal Reserve Bank of New York, 1965).

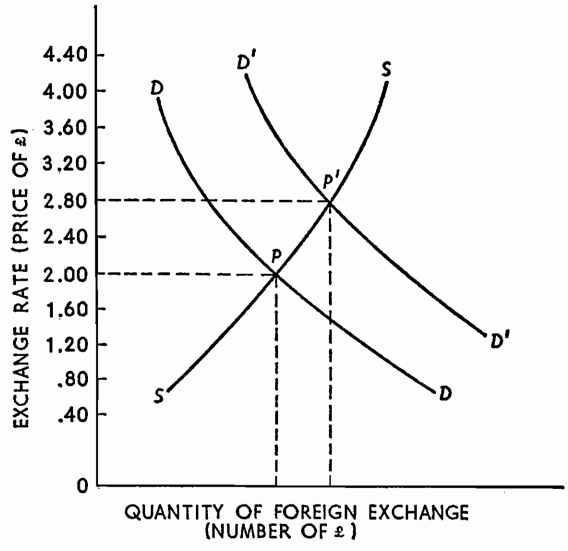
these currencies are little used in international trade. Instead, international payments involving these countries are usually made in dollars or, to a lesser extent, in sterling. For this reason, the dollar and sterling are used much more extensively than other currencies in the making of private international payments.

The Foreign Exchange Market: A Two-Country Model

An exchange rate is simply a price of one currency in terms of another; in a free market, exchange rates would be determined by the supply of and demand for foreign exchange. To illustrate, let us suppose there are only two

FIGURE 21–1

Supply of and Demand for Foreign Exchange



countries, the United States and the United Kingdom, and that all payments between the two countries are made in pounds. Then the market for pounds in the United States might be as depicted in Figure 21–1.[3] Initially, *DD* is

[3] If some payments were made in dollars, there would also be a market for dollars in the United Kingdom, and the supply and demand for dollars would determine an exchange rate in this market. However, unless this exchange rate was the same as the rate determined in the market for pounds in the United States, arbitrage would occur between the two markets to bring the rates to equality. For example, if the exchange rate was $2.40 in New York and $2.42 in London, arbitragers would buy pounds with dollars in New York and simultaneously sell them for dollars in London, thereby realizing a profit of 2 cents to the pound. Such arbitrage would add to the demand for pounds and raise the exchange rate in New York and would add to the supply of pounds and depress the rate in London. This process would continue until the same rate prevailed in both markets.

the demand for pounds by Americans for use in paying for purchases of goods, services, and securities from the United Kingdom. *SS* is the supply of pounds derived from U.S. exports of goods, services and securities to the United Kingdom. The market is equilibrated at point *P* with an exchange rate of £1 = $2. The demand curve *DD* slopes downward to the right, because a fall in the exchange rate makes pounds (and therefore British goods, services, and securities) less expensive in the United States, inducing Americans to purchase larger quantities of them. The supply curve *SS,* on the other hand, slopes upward to the right, because a rise in the exchange rate means that a pound will command more dollars, thereby making American goods, services, and securities cheaper in the United Kingdom and causing more pounds to be spent on them.[4] Changes in demand or supply would cause the exchange rate to adjust in order to maintain equilibrium in the market. For example, if the demand for pounds should shift to *D′D′* in Figure 21–1, the new market equilibrium would be at *P′,* with an exchange rate of $2.80.

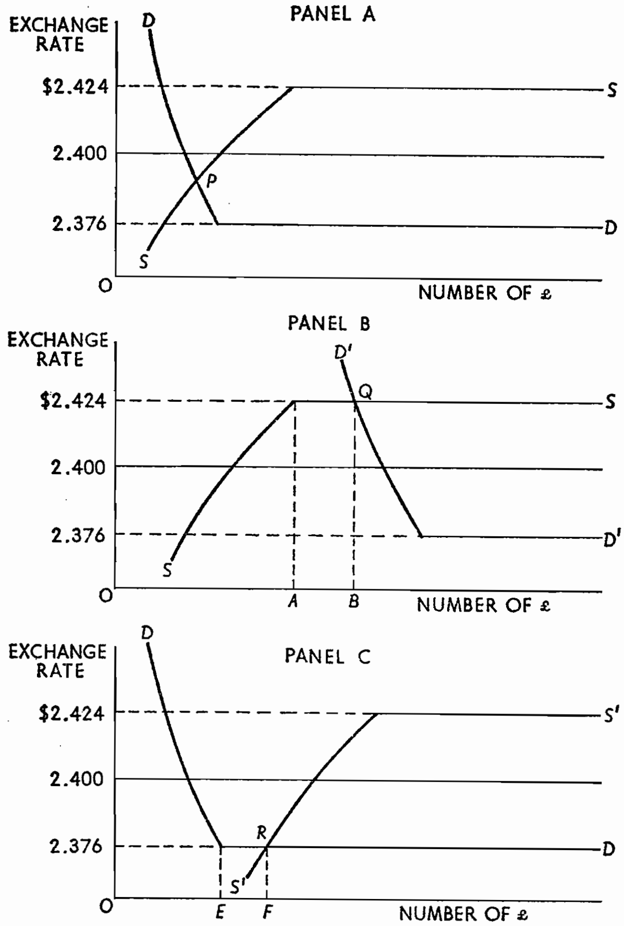
Stabilization of Exchange Rates

Under the present international monetary system, foreign exchange markets are relatively free of controls in the advanced countries outside the communist bloc, including the United States, Canada, Japan, the United Kingdom, and Western Europe. However, the IMF requires its members to take measures to prevent exchange rates from departing from the established par values by more than 1 percent in either direction.

The way such an arrangement would work in the two-country case depicted in Figure 21–1 can be explained with the aid of Figure 21–2. Suppose the United Kingdom takes responsibility for keeping the pound-dollar exchange rate within 1 percent of the $2.40 parity. It does this by buying and selling dollars in the foreign-exchange market. The upper bound to the exchange rate is $2.424 ($2.40 × 1.01). The Bank of England, acting on behalf of the United Kingdom, stands ready to buy dollars with pounds—that is, supply pounds to the market—in whatever quantity is demanded at $2.424 per pound. Thus, the supply of pounds becomes perfectly elastic at that price; the supply curve is depicted by the line *SS* in Panel A of Figure 21–2. The rising portion of the curve represents the

[4] Actually a rise in the exchange rate will clearly induce British buyers to purchase more dollars and more American goods, but it is not certain that they will spend more pounds on these purchases. This depends on the elasticity of British demand for American goods, as will be shown in more detail later in Chapter 23. This is, however, a minor point that can be disregarded in the context of the present discussion. For an extraordinarily lucid and complete discussion of the foreign exchange markets, the reader is referred to Fritz Machlup, "The Theory of Foreign Exchanges," *Economica,* Vol. VI (New Series), November 1939, pp. 375–97 and February 1940, pp. 23–49, reprinted in H. S. Ellis and L. A. Metzler (eds.), *Readings in the Theory of International Trade* (Philadelphia: Blakiston Co., 1949), pp. 104–58.

FIGURE 21–2
Stabilization of Exchange Rates



supply of pounds entering the market as a result of U.S. exports of goods, services, and securities to the United Kingdom, while the horizontal portion, at an exchange rate of \$2.424, reflects the willingness of the United Kingdom to supply an unlimited quantity of pounds at that price. The lower bound to the exchange rate is \$2.376 (\$2.40 × 0.99). The Bank of England stands ready to sell dollars for pounds—that is, absorb pounds from the market—in whatever quantity is offered at \$2.376 per pound. Thus the demand for pounds becomes perfectly elastic at that price; the demand curve is depicted by the line *DD* in Panel A of Figure 21–2. The falling portion of the curve represents the demand for pounds in the market as a result of U.S. imports of goods, services, and securities from the United Kingdom, while the horizontal portion at an exchange rate of $2.376 reflects the willingness of the United Kingdom to absorb an unlimited quantity of pounds at that price.

In the case depicted in Panel A of Figure 21–2, the demand for and supply of foreign exchange are equated at point *P* where the exchange rate is $2.39, which lies between the upper and lower bounds of fluctuation permitted. The supply of pounds entering the market as a result of U.S. exports to the United Kingdom is equal to the demand for pounds to pay for U.S. imports from the United Kingdom, and no intervention in the foreign exchange market by the Bank of England is necessary.

Now suppose that the demand for pounds by Americans to pay for imports of goods, services, and securities from the United Kingdom increases, so that the demand curve shifts from *DD* in Panel A to *D′D′* in Panel B. With the supply curve the same as before, the demand for and supply of pounds now intersect at point *Q* in Panel B. The exchange rate rises to its upper bound of $2.424. The number of pounds demanded by Americans to pay for imports from the United Kingdom is *OB,* while the number of pounds entering the market as proceeds from U.S. exports to the United Kingdom is only *OA.* To fill the gap, the Bank of England must supply *AB* of pounds by purchasing dollars in the foreign exchange market.

Suppose, on the other hand, that the demand curve for pounds remains in the initial position depicted in Panel A of Figure 21–2, but that the supply curve shifts to the position *S′S′* as shown in Panel C as a result of an increase in the U.K. demand for U.S. goods, services, and securities. In this case, the demand and supply curves intersect at point *R* in Panel C, and the exchange rate falls to its lower bound of $2.376. The number of pounds coming into the market as proceeds of U.S. exports to the United Kingdom is *OF,* while the number of pounds demanded to pay for U.S. imports from the United Kingdom is only *OE.* In order to keep the exchange rate from falling below $2.376, it is necessary for the Bank of England to enter the market and buy the excess pounds in the amount *EF,* making payment with dollars.

In the case depicted in Panel A, trade in goods, services, and securities between the two countries is balanced in the sense that proceeds from U.S. exports are just sufficient to cover payment for U.K. imports, and the same is true for the United States. Panel B depicts a case in which the United Kingdom has a balance-of-payments surplus—that is, proceeds from U.K. exports are more than sufficient to cover U.K. imports, and the balance accrues to the Bank of England in the form of increased holdings of dollars. On the other hand, in the case depicted in Panel C, the United Kingdom has a balance-of-payments deficit—payments for U.K. imports exceed receipts from U.K. exports, and the balance is settled by an outflow of dollars from the Bank of England. Conversely, of course, the United States has a balance-of-payments deficit in Panel B and a balance-of-payments surplus in Panel C.

The Need for Reserves

In order to be able to absorb the excess pounds from the market when it experiences a deficit, as in Panel C, it is necessary for the United Kingdom to have "reserves" in the form of dollars to pay for the pounds or to have some means of acquiring the necessary dollars. In general, these reserves must have been built up in the past as a result of balance-of-payments surpluses of the sort depicted in Panel B.

The United Kingdom meets its obligation to maintain exchange stability, as required by the IMF, by standing ready to buy dollars at an exchange rate of $2.424 and to sell dollars at an exchange rate of $2.376. Thus, it keeps the exchange rate from moving outside this range.[5] The United States, on the other hand, satisfies its obligation to the IMF by standing ready to buy gold from the United Kingdom, or sell gold to that country in unlimited quantities at $35 an ounce. This serves to establish a fixed link between both currencies and gold.

When the United Kingdom has a balance-of-payments surplus (as in Panel B of Figure 21–2), its reserves will increase, and it has a choice as to the form in which it will hold these reserves: it can either hold the dollars it acquires from the market or it can use the dollars to acquire gold from the United States at the fixed price of $35 an ounce. If it chooses to hold the dollars—which initially accrue to it in the form of deposits in U.S. banks—it can invest these dollars in interest-bearing securities in the United States, thereby earning some return on its reserves. However, while currency parities are fixed in terms of gold at any particular time under the IMF, there are provisions, to be discussed below, for changing these parities in certain circumstances. If, for example, the United States should experience a persistent deficit in its balance of payments which it wanted to eliminate, it might raise the price of gold, which (as will be explained in detail in Chapter 23) would make pounds dearer in terms of dollars and dollars cheaper in terms of pounds, thereby stimulating U.S. exports and discouraging U.S. imports. The possibility that the United States might take action of this kind creates a risk for the United Kingdom in holding its reserves in dollars rather than in gold. That is, if the United Kingdom converted dollars into gold at the existing price of $35 an ounce and the United States should then raise the price of gold to, say, $42, the United Kingdom could reconvert the gold into dollars at the latter price when it needed additional dollars.

---

[5] In practice, the Bank of England will intervene in the market by buying or selling dollars before the exchange rate reaches these limits. The main reason for this is a desire to prevent speculation in anticipation of a downward or upward revision of the parity value of sterling. Such speculation will be further discussed in Chapter 24.

In this system, the dollar is a "reserve currency"; the United Kingdom may hold its reserves in either dollars or gold—or, more likely, a combination of the two. The United States holds its reserves in gold. It may be noted that reserves of the world as a whole—that is, of the United States and United Kingdom together—are increased as a result of additional gold entering monetary reserves, either from new gold production or through the release of existing gold from nonmonetary uses, and also when the United States has balance-of-payments deficits that are settled through acceptance of additional dollars by the United Kingdom.

The Present System

Although the present international monetary system is considerably more complicated than the example just presented because there are many countries rather than just two, its essential features are the same as in our example. Currencies of all member countries of the IMF are defined in gold, and this establishes a set of exchange parities among them. Countries other than the United States keep their dollar exchange rates within a range of not more than 1 percent either side of parity by buying and selling dollars in exchange for their own currencies.[6] The United States in turn buys and sells gold freely at $35 an ounce.[7] Thus, countries other than the United States take responsibility for maintaining exchange stability between their currencies and the dollar, while the United States takes responsibility for maintaining a fixed link between the dollar and gold—and therefore indirectly between all other currencies and gold.

Most member countries of the IMF hold their international monetary reserves in gold and dollars, with the ratio of dollars to gold varying substantially among countries. One slight complicating factor is that the countries of the Sterling Area keep their currencies linked to the British pound sterling and hold their reserves in sterling.[8] Britain in turn keeps the pound tied to the dollar in the same manner as most other countries outside the Sterling Area.

The Role of the IMF

Under the present international monetary system, countries other than the United States hold reserves in the form of foreign currencies—primarily

[6] In practice, most countries do not permit their dollar exchange rates to fluctuate by the full 1 percent permitted by the IMF rules. See footnote 5 above.

[7] Actually, the United States maintains a small margin between its buying and selling prices. It sells gold at $35.0875 per ounce, one fourth of 1 percent above the $35 parity, and buys it at $34.9125 per ounce, one fourth of 1 percent below parity.

[8] Members of the Sterling Area include Australia, India, Malaysia, New Zealand, Nigeria, Pakistan, South Africa, and a number of smaller countries in Asia, Africa, the South Seas, and the Caribbean having an historic association with Britain.

dollars—or in gold that they can turn into dollars by sale to the U.S. Treasury. These reserves are needed in order to permit a country to purchase the excess supply of its currency in the foreign exchange market to prevent the exchange rate from falling more than 1 percent below the IMF parity when its balance of payments goes into deficit—in circumstances analogous to those experienced by the United Kingdom in Panel C of Figure 21–2. The United States holds reserves in gold—and, to a limited extent, in foreign currencies—in order to be able to redeem dollars that are presented to it for conversion by other countries.

One of the major functions of the IMF is to serve as a source from which member countries experiencing balance-of-payments deficits can borrow foreign currencies to supplement their holdings of gold and foreign exchange reserves. Each member country is assigned a "quota" and is required to pay into the Fund an amount equal to its quota, 75 percent of the payment normally to be made in the country's own currency and the remaining 25 percent in gold.[9] Through the payment into the Fund of an amount equal to its quota by each country, the Fund began its operations with a stock of national currencies and gold. Member countries receive "drawing rights" against the Fund that are related to their quotas. If a member country needs reserves to cover a balance-of-payments deficit it can draw foreign currencies from the Fund, paying in an equivalent amount of its own currency in exchange. That is, it can, in effect, borrow foreign currencies from the Fund. Such borrowings are to be used to deal with temporary balance-of-payments deficits and are normally to be repaid within a period of three to five years. The borrowing country may repay by "repurchasing" its currency from the Fund, paying in convertible currencies or gold in exchange. Alternatively, borrowings are effectively repaid if another country, experiencing a balance-of-payments deficit, draws the borrowing country's currency from the Fund.

A country is permitted to borrow from the Fund, virtually on notice, an amount equal to the difference between its quota and the Fund's holdings of its currency. This amount is termed the country's "reserve position in the IMF" and is counted as part of the country's monetary reserves. Initially, a country's reserve position normally equals its gold contribution—its so-called "gold tranche," equal to 25 percent of its quota. However, if other countries have drawn its currency from the Fund, this enlarges the amount the country can borrow freely. For example, on June 30, 1969, the Fund held only $68.5 million of the $112.5 million of Norwegian kroner that was initially paid in by Norway; hence, Norway, which has a quota of $150 million and a "gold tranche" of $37.5 million, would, in principle, have

---

[9] The Fund's holdings of the currencies of its members are in the form either of deposits in the country's central bank or of noninterest-bearing securities that are redeemable on demand when the Fund needs to use the country's currency.

been able to draw without restriction an amount equal to $81.5 million ($150 million minus $68.5 million). The reason for this is that other countries that had had deficits had drawn $44 million in Norwegian kroner from the Fund. These drawings, in effect, established a right by Norway to draw an equivalent amount from the Fund.

The upper limit of the total amount of borrowing that a country may have outstanding from the Fund at any time is reached when the Fund's holdings of its currency reach 200 percent of its quota. Thus, the total amount that Norway could, in principle, have borrowed from the Fund as of June 30, 1969, was $231.5 million ($2 \times 150 - 68.5 = 231.5$). However, when the Fund's holdings of a country's currency become equal to its quota, the Fund may require it to satisfy certain conditions to qualify for further borrowing; in particular, it may insist that the country take policy action to reduce or terminate the balance-of-payments deficit that is the cause of the borrowing. Indeed, the conditions imposed on the country may become more and more stringent as its borrowings increase. In addition, the interest rate the Fund charges for borrowings becomes higher as the size and duration of a country's borrowings increase.

THE BALANCE OF PAYMENTS

The balance of payments is a summary of all the international transactions of a country and its citizens during a specified period of time. Table 21–1 presents the balance of payments of the United States for 1966. The balance of payments is a double-entry accounting statement, and, as is the case with any such statement, every transaction is entered twice, once as a debit and once as a credit. In general, a credit, which is entered in the balance of payments with a positive sign, represents the export of something, whether it be a commodity, a service of some kind, a security, a bank deposit, or gold. Likewise, generally speaking, a debit, which is entered with a negative sign, represents an import of something, whether it be commodity, service, security, deposit, or gold.

Thus, if an import of goods into the United States from the United Kingdom is paid for by the transfer of a bank deposit from the United States to the United Kingdom, the import of goods constitutes the debit side of the transaction in the U.S. balance of payments, and the transfer of the deposit constitutes the credit side. In effect, what has happened is that the United States has imported goods and exported a deposit claim against its banking system. Similarly if American investors receive interest or dividends on investments in European stocks or bonds, the credit is to income on investments in the U.S. balance of payments, which represents an export of the services of U.S. capital, while the transfer of deposits constitutes the debit side of the transaction. The United States has exported capital services and imported deposit claims.

The U.S. Balance of Payments

Table 21–1 shows the net debits or net credits in the U.S. balance of payments, classified under a number of headings. Thus, under the first heading it shows that the United States had a credit balance of $2,516 million on goods and services account in 1968—that is, that we exported $2,516 million more goods and services than we imported. The gross transactions in goods and services were, of course, much larger than this; actually, our exports of goods and services were $50,594 million while our

TABLE 21–1

Balance of Payments of the United States, 1968
(minus sign indicates net payments (debits); absence of sign indicates net receipts; amounts in millions of dollars)

Balance on goods and services		2,516
Merchandise	626	
Military expenditures	−3,103	
Transportation	−324	
Travel	−1,252	
Income on investments	4,766	
Other services	1,803	
Remittances and pensions		−1,159
U.S. government grants and loans		−3,955
Long-term capital movements		−506
Short-term capital movements, exclusive of changes in U.S. liquid liabilities to foreign central banks and governments		5,384
Errors and unrecorded transactions		−642
Surplus or deficit (−), official reserve transactions basis		1,638
Financing of deficit or surplus:		
U.S. official reserve assets (increase, −)		−880
Gold	1,173	
Convertible currencies	−1,183	
IMF reserve position	−870	
Liabilities to foreign official agencies (decrease, −)		−758

SOURCE: Department of Commerce, *Survey of Current Business*, September 1969, pp. 36–38.

imports were $48,078 million. The subdivisions under the balance on goods and services show the breakdown by categories; thus, our exports of merchandise exceeded our imports by $626 million, our purchases of transportation services from other countries exceeded our sales of such services to them by $324 million, and so on.

The category, "military expenditures" includes amounts spent abroad by U.S. military personnel, as well as our foreign expenditures for military construction and services and cash expenditures abroad for military assistance to our allies. However, military grants in the form of goods and services, which were estimated to be $838 million in 1968, are not included in the balance of payments as presented in Table 21–1. If they were included, they would produce a credit in the form of an increase in exports of goods and services and an equal debit in the form of an increase in the category "U.S. government grants and loans," to be discussed below. Since inclusion of these military shipments in exports would exaggerate the size of our credit balance on goods and services account, it seems preferable to omit them.

Remittances and pensions include gifts made in the form of goods or money by individuals and by charitable or other institutions, together with pension payments to persons residing overseas. The debit balance of $1,159 million shown in Table 21–1 indicates that such transfers out of the United States exceeded transfers into the United States by this amount.

The category, "U.S. government grants and loans," represents roughly the U.S. foreign economic aid program; as indicated above, military grants are excluded. For balance of payments purposes, government grants are defined as transfers of resources—whether in money or in kind—for which no payment is expected, or for which repayment terms have not yet been defined. Loans include all U.S. government loans, repayments of such loans, investments in international institutions, and net changes in other U.S. government claims except those included in official reserves, to be discussed below. The debit balance in Table 21–1 indicates a net outflow of $3,955 million from the United States in 1968.

In the capital account of the balance of payments, foreign lending by U.S. citizens, business firms, or financial institutions—i.e., the importing of securities or I O U's—is entered as a debit, while U.S. borrowing from abroad—the exporting of securities or I O U's—is entered as a credit. It is customary to distinguish between long-term and short-term capital movements. Somewhat arbitrarily, long-term capital movements are defined to include the export and import of securities or loan contracts having a maturity of more than one year at the time of original issuance, while short-term capital movements are those involving export or import of securities or loan contracts having an original maturity of one year or less.

Table 21–1 shows that the United States experienced a long-term private capital outflow of $506 million in 1968—that is, that we increased our long-term private claims by that amount. This is a net figure, and it includes the changes in both U.S. holdings of foreign long-term claims and foreign holdings of long-term claims against the United States. On short-term capital account, there was a net inflow of $5,384 million. This includes both changes in U.S. holdings of short-term claims against foreigners and changes in foreign holdings of short-term claims against the United States. It does not, however, include holdings by foreign central banks and governments of claims against the United States.

Many of the items in the balance of payments are estimated from incomplete data and are subject to errors of uncertain magnitude. In many cases, the debit or credit side of a transaction may appear in the records while the other side may be missed. If all transactions were fully recorded, debits should equal credits for the balance of payments as a whole. But because errors are made and portions of some transactions are missed, this is never the case. To make the accounts balance as they should, the difference between the total recorded debits and the total recorded credits is entered under the heading "errors and unrecorded transactions." In 1968, as it turned out, total recorded credits exceeded total recorded debits by $642 million, and this amount was therefore shown as a debit under "errors and unrecorded transactions." There is no way, of course, of knowing how this amount should be allocated among the various items in the balance of payments, although there are some reasons for suspecting that a large portion of it is frequently unrecorded capital movements.

Definition of Surplus or Deficit

We come now to the problem of defining the surplus or deficit in the balance of payments. Since the balance of payments is a double-entry accounting statement, the sum of all the debits must necessarily equal the sum of all the credits, that is, the statement in total must necessarily balance. Consequently, when it is said that there is a surplus—receipts exceed payments—or a deficit—payments exceed receipts—it must be that only a portion of the balance of payments is being considered. Of course, the surplus or deficit could be defined in various ways; that is, various categories of items could be used in calculating it. However, we can get a clue concerning the appropriate definition under the present international monetary system by referring back to Figure 21–2 and the related discussion. We noted that in the case depicted in Panel B of Figure 21–2, the number of pounds demanded by Americans to pay for imports from England was *OB*, while the number of pounds entering the market as proceeds from U.S. exports to the United Kingdom was only *OA*. To keep the demand from exceeding the supply and forcing the exchange rate above the upper limit of $2.424, it was necessary for the Bank of England to supply *AB* of pounds by purchasing dollars in the foreign exchange market. Thus, the reserves held by the Bank of England increased by the amount *AB*. The Bank of England has the option of holding these dollars as part of its reserves (in the form of deposits in U.S. banks or securities purchased in the United States) or turning them in to the U.S. Treasury for gold at the fixed price of $35 an ounce. It would be logical to say that the United Kingdom had a surplus in its balance of payments in the amount *AB*, that is, to define the surplus to be equal to the increase in its reserves. Alternatively, the surplus is equal to the amount by which U.K. exports of goods, services, and securities to the United States (the distance *OB*) exceed U.K. imports of goods, services, and securities from the United States, not counting the purchases of dollars by the Bank of England (the distance *OA*).

The opposite case is depicted in Panel C of Figure 21–2. Here the supply of pounds entering the foreign exchange market as a result of U.K. imports of goods, services, and securities exceeds the demand for pounds to pay for U.S. imports from the United Kingdom. In order to prevent the supply of pounds from exceeding the demand and forcing the exchange rate below the lower limit of $2.376, it was necessary for the Bank of England to buy *EF* of pounds by supplying dollars to the foreign exchange market. Thus, the reserves held by the Bank of England declined by the amount *EF*. In this case, it would be logical to say that the United Kingdom had a deficit in its balance of payments in the amount *EF*—i.e., to define the deficit to be equal to the decline in U.K. reserves. Alternatively, the deficit is equal to the amount by which U.K. imports of goods, services, and securities (the distance *OF*) exceeds U.K. exports of goods, services, and securities, not counting the sales of dollars by the Bank of England (the distance *OE*).

Thus, in the case of a nonreserve currency country, it would be appropriate to define the surplus or deficit in the balance of payments to be equal to the increase or decrease in the country's reserves. In the case of a reserve currency country such as the United States, the situation is a little more complicated. When the United States has a deficit, in the sense that the outflow of dollars exceeds the inflow, U.S. reserves of gold and foreign exchange need not necessarily decline. Foreign central banks, fulfilling their responsibilities for stabilizing exchange rates under the IMF rules, will be forced to buy dollars (as the United Kingdom was forced to do in the case depicted in Panel B of Figure 21–2). It is quite likely, however, that the central banks will decide to hold some portion of the dollars they have acquired in the form of deposits in U.S. banks or investments in securities purchased in the United States. To the extent that they decide to do this—i.e., to continue to hold the increase in their reserves in the form of dollars—U.S. reserves are not reduced when the United States has a deficit. On the other hand, to the extent that they choose to turn the dollars in to the U.S. Treasury for gold at the price of $35 an ounce, U.S. reserves are reduced. Conversely, if the United States should have a balance-of-payments surplus, so that foreign central banks would have to sell dollars to stabilize exchange rates (as in Panel C of Figure 21–2), this would not increase U.S. reserves of gold and foreign exchange except to the extent that foreign central banks sold gold to the U.S. Treasury to replenish the losses of dollars resulting from their sales in the foreign exchange market.

Thus, while the surplus or deficit of a nonreserve currency country may be defined to be equal to the increase or decrease of its reserves, a modification of this way of measuring the deficit is required in the case of a reserve currency country. For the United States, the balance-of-payments surplus or deficit may be defined to be equal to the algebraic sum of the increase or decrease in reserves plus the decrease or increase in holdings of dollars by foreign central banks. This is the way in which the deficit or surplus was defined in the construction of Table 21–1.

The student's attention is directed to the line drawn across Table 21–1 beneath the entry for "errors and unrecorded transactions." The categories of transactions recorded above this line are those responsible for the deficit: trade in goods and services, remittances and pensions, U.S. government grants and loans, long-term capital movements, short-term capital movements, and errors and unrecorded transactions. In 1968, the combined effect of these categories of transactions was to produce a net inpayment or credit balance of $1,638 million. This is entered just below the line as the balance-of-payments surplus under the specific designation "surplus or deficit (—), official reserve transactions basis." As indicated above, the surplus or deficit is equal to the algebraic sum of the increase or decrease in reserves plus the decrease or increase in liabilities to foreign central banks. The changes in reserves and liabilities accounting for the deficit are presented under the heading, "financing of deficit or surplus." In 1968, U.S. reserves, including gold, convertible currencies, and reserve position in the IMF, increased by $880 million.[10] At the same time, U.S. liabilities to foreign official agencies (i.e., foreign central banks and monetary authorities) declined by $758 million. The increase in reserves plus the decline in liabilities amounted to $1,638 million, accounting for the surplus of that amount.[11]

[10] The increase of $870 million in the U.S. reserve position in the IMF is a result of transactions which reduced IMF holdings of dollars by that amount. These transactions can be summarized as follows: (1) The United States repurchased dollars from the IMF with foreign currencies (i.e., repaid its earlier borrowings from the IMF) in the amount of $84 million; (2) other countries drew (i.e., borrowed) dollars from the IMF in the amount of $806 million; and (3) the IMF earned income in dollars to the extent of $20 million. See the discussion of the reserve position in the IMF, pp. 451–52. above.

[11] The *official reserve transactions* surplus or deficit which is used in Table 21–1 and discussed above is one of two surplus or deficit concepts used by the Department of Commerce in official balance-of-payments tabulations. The other, known as the *liquidity* surplus or deficit, differs in one important respect: it excludes the change in foreign private holdings of liquid claims against the United States in calculating the surplus or deficit itself and instead treats this item as a means of financing the surplus or deficit. In other words, this item, which is included in "short-term capital movements" above the line in Table 21–1, is transferred below the line in drawing up the balance of payments on a liquidity basis. Prior to December 1965, the liquidity concept was the only one used. However, in its report submitted in 1965, the Review Committee for Balance of Payments Statistics recommended the official reserve transactions concept as the best measure of the surplus or deficit to be used in official reports. For a discussion of the relative merits of the two concepts, see *The Balance of Payments of the United States: A Review and Appraisal,* Report of the Review Committee for Balance of Payments Statistics to the Bureau of the Budget (Washington: U.S. Government Printing Office, 1965), chap. 9. The present writer prefers the official reserve transactions concept, and it fits much more logically into the discussion of the foreign exchange markets than does the other.

Chapter 22 RELATIONS BETWEEN DOMESTIC ECONOMIC ACTIVITY AND INTERNATIONAL TRADE

In a country that maintains trade relations with other countries, both the domestic economy and the balance of payments are affected by developments related to foreign trade and by economic conditions in other countries. The relationships involved are rather complex, but we shall attempt in this chapter to sort out the main forces at work. For simplicity, we will first consider changes in national income in the absence of changes in prices; later we will consider the effects of price changes.

NATIONAL INCOME AND INTERNATIONAL TRADE[1]

Let us consider a world in which there are two countries, which we shall call Country 1 and Country 2. For some purposes, it will be useful to think of Country 1 as the country we are primarily concerned with (say, the United States) and Country 2 as the rest of the world. The economies of these two countries are governed by the following set of equations:

$$C_1 = (1 - s_1)Y_1 + \bar{C}_1 \quad (1)$$
$$M_1 = m_1Y_1 + \bar{M}_1 \quad (2)$$
$$Y_1 = C_1 + I_1 + X_1 - M_1 \quad (3)$$
$$C_2 = (1 - s_2)Y_2 + \bar{C}_2 \quad (4)$$
$$M_2 = m_2Y_2 + \bar{M}_2 \quad (5)$$
$$Y_2 = C_2 + I_2 + X_2 - M_2 \quad (6)$$
$$X_1 = M_2 \quad (7)$$
$$X_2 = M_1 \quad (8)$$

In these equations, subscript 1 refers to Country 1 and subscript 2 to Country 2; Y is national income (GNP), C is personal consumption expenditures, I is gross private domestic investment, X is exports of goods

[1] For more extensive treatments of this subject, see Fritz Machlup, *International Trade and the National Income Multiplier* (Philadelphia: Blakiston Co., 1943); Lloyd A. Metzler, "Underemployment Equilibrium and International Trade," *Econometrica*, Vol. X, April 1942, pp. 97–112; and Romney Robinson, "A Graphical Analysis of the Foreign Trade Multiplier," *Economic Journal*, Vol. LXII, September 1952, pp. 546–64.

and services, and M is imports of goods and services. As is the case in the national income accounts, consumption and investment include expenditures on imported as well as domestically produced goods and services; this is the reason for the subtraction of imports (M_1 and M_2 in Equations 3 and 6).[2] Equations 1 and 4 are the consumption functions of Countries 1 and 2, respectively. The parameters s_1 and s_2 are the respective marginal propensities to save; that is, $1 - s_1$ and $1 - s_2$ are the marginal propensities to consume. Equations 2 and 5 are the import demand equations for Country 1 and Country 2, respectively. It is assumed that the demand for imports in each country depends on that country's income; the parameters m_1 and m_2 are referred to as the "marginal propensities to import" of the two countries.[3] Equations 7 and 8 merely state the fact that in a two-country system, the exports of one country are necessarily equal to the imports of the other.

In using this system of equations, we are making several assumptions, partly to simplify the presentation and partly to isolate the effects of changes in income from other changes that might be occurring at the same time.

1. We are disregarding the effects of government expenditures and taxes, assuming, in effect, that there is no government. This assumption is made entirely for simplicity; the analysis could quite easily be modified to take explicit account of government activities.

2. We are assuming that the exchange rate between the currencies of the two countries is rigidly fixed and also that internal prices do not change in either country. In other words, we assume that in both countries there are unemployed productive factors, constant returns in production, and constant factor prices. Real consumption and real imports are taken to be functions of real income. With constant prices and exchange rates, real magnitudes are proportional to money magnitudes; hence money consumption and money imports are functions of money income.

3. We are disregarding all monetary influences, assuming, in effect, that interest rates remain unchanged in both countries.

4. Possible effects of the level of national income or of changes in national income (the acceleration principle) on domestic investment are neglected.

5. No account is taken of elements in the balance of payments other than trade in goods and services.

In this model, investment in each country (I_1 and I_2) is taken to be exogenously determined. The parameters, $\bar{M}_1$, $\bar{M}_2$, $\bar{C}_1$, and $\bar{C}_2$ reflects the level of import and consumption demand. If I_1, I_2, $\bar{M}_1$, $\bar{M}_2$, $\bar{C}_1$, $\bar{C}_2$, s_1, s_2, m_1, and m_2 are all taken as given, the model contains eight equations, which

---

[2] See Chapter 2, pp. 29–30.

[3] There is a good deal of empirical evidence to support the assumption that import demand is a function of income. Of course, it is also a function of other variables, particularly prices of imports relative to domestically produced goods and services, which are assumed not to be operative in the above model.

are sufficient to determine the eight variables, Y_1, Y_2, C_1, C_2, M_1, M_2, X_1, and X_2.

FOREIGN TRADE MULTIPLIERS

From this model we will derive a set of *foreign trade multipliers* which show the effects on income in the two countries of autonomous changes in investment (I_1 and I_2) and of autonomous shifts in import demand (changes in $\bar{M}_1$ and $\bar{M}_2$). In other words, we will extend the concept of the multiplier developed in Chapter 6 to an economy which carries on trade with another economy. The easiest way to begin the derivation of the foreign trade multipliers is to write Equations 1 to 8 in terms of changes:

$$\Delta C_1 = (1 - s_1)\Delta Y_1 + \Delta\bar{C}_1 \tag{9}$$
$$\Delta M_1 = m_1 \Delta Y_1 + \Delta\bar{M}_1 \tag{10}$$
$$\Delta Y_1 = \Delta C_1 + \Delta I_1 + \Delta X_1 - \Delta M_1 \tag{11}$$
$$\Delta C_2 = (1 - s_2)\Delta Y_2 + \Delta\bar{C}_2 \tag{12}$$
$$\Delta M_2 = m_2 \Delta Y_2 + \Delta\bar{M}_2 \tag{13}$$
$$\Delta Y_2 = \Delta C_2 + \Delta I_2 + \Delta X_2 - \Delta M_2 \tag{14}$$
$$\Delta X_1 = \Delta M_2 \tag{15}$$
$$\Delta X_2 = \Delta M_1 \tag{16}$$

Since the equations are linear and the slopes ($1 - s_1$, $1 - s_2$, m_1, and m_2) are assumed not to change, the equations can be written in this way. Thus, Equation 9 says, "The change in consumption in Country 1 is equal to the marginal propensity to consume in Country 1 times the change in income in Country 1 plus the autonomous change in consumption in Country 1."

We can use Equations 15 and 16 as a basis for substituting ΔM_2 for ΔX_1 in Equation 11 and ΔM_1 for ΔX_2 in Equation 14. Then Equations 11 and 14 become

$$\Delta Y_1 = \Delta C_1 + \Delta I_1 + \Delta M_2 - \Delta M_1$$
$$\Delta Y_2 = \Delta C_2 + \Delta I_2 + \Delta M_1 - \Delta M_2$$

Substituting in these equations the values of ΔC_1, ΔM_1, ΔC_2, and ΔM_2 given by Equations 9, 10, 12, and 13, we obtain, after some simplification,

$$(s_1 + m_1)\Delta Y_1 = \Delta\bar{C}_1 + \Delta I_1 - \Delta\bar{M}_1 + m_2 \Delta Y_2 + \Delta\bar{M}_2 \tag{17}$$
$$(s_2 + m_2)\Delta Y_2 = \Delta\bar{C}_2 + \Delta I_2 - \Delta\bar{M}_2 + m_1 \Delta Y_1 + \Delta\bar{M}_1 \tag{18}$$

From these equations, we can readily derive the multipliers we are interested in.

Case 1: Multipliers Applicable to an Autonomous Change in Domestic Investment in Country 1 (ΔI_1). In this case, everything in Equations 17 and 18 except investment in Country 1, income in Country 1, and income in Country 2 remains constant. Thus, $\Delta\bar{C}_1 = 0$, $\Delta\bar{M}_1 = 0$, $\Delta\bar{M}_2 = 0$, $\Delta\bar{C}_2 = 0$, and $\Delta I_2 = 0$. Equations 17 and 18 become

$$(s_1 + m_1)\,\Delta Y_1 = \Delta I_1 + m_2\,\Delta Y_2$$
$$(s_2 + m_2)\,\Delta Y_2 = m_1\,\Delta Y_1$$

Solving the second of these equations for ΔY_2, we obtain

$$\Delta Y_2 = \frac{m_1}{s_2 + m_2}\,\Delta Y_1 \tag{19}$$

Substituting this into the first of the equations, we obtain,

$$(s_1 + m_1)\,\Delta Y_1 = \Delta I_1 + \frac{m_1 m_2}{s_2 + m_2}\,\Delta Y_1$$

Solving for ΔY_1, we have

$$\Delta Y_1 = \frac{s_2 + m_2}{s_1 s_2 + s_1 m_2 + m_1 s_2}\,\Delta I_1 \tag{20}$$

Substituting this value for ΔY_1 in Equation 19 and solving for ΔY_2, we obtain

$$\Delta Y_2 = \frac{m_1}{s_1 s_2 + s_1 m_2 + m_1 s_2}\,\Delta I_1 \tag{21}$$

The trade balance of Country 1 (B_1) is, by definition, the difference between Country 1's exports and its imports—that is $B_1 = X_1 - M_1$. Thus, the *change* in its trade balance is

$$\Delta B_1 = \Delta X_1 - \Delta M_1 = \Delta M_2 - \Delta M_1$$
$$\Delta B_1 = m_2\,\Delta Y_2 - m_1\,\Delta Y_1$$

Substituting the values of ΔY_1 and ΔY_2 from Equations 20 and 21 and simplifying, we have

$$\Delta B_1 = \frac{-m_1 s_2}{s_1 s_2 + s_1 m_2 + m_1 s_2}\,\Delta I_1 \tag{22}$$

(Of course, the change in the trade balance of Country 2, ΔB_2, is the same as the change in the trade balance of Country 1, but with opposite sign.)

To illustrate these results, suppose $s_1 = 0.45$, $m_1 = 0.05$, $s_2 = 0.22$, and $m_2 = 0.20$. Using Equations 20, 21, and 22, this yields

$$\frac{\Delta Y_1}{\Delta I_1} = \frac{0.22 + 0.20}{(0.45)(0.22) + (0.45)(0.20) + (0.05)(0.22)} = 2.1$$
$$\frac{\Delta Y_2}{\Delta I_1} = \frac{0.05}{(0.45)(0.22) + (0.45)(0.20) + (0.05)(0.22)} = 0.25$$
$$\frac{\Delta B_1}{\Delta I_1} = \frac{-(0.05)(0.22)}{(0.45)(0.22) + (0.45)(0.20) + (0.05)(0.22)} = -0.055$$

Thus, an autonomous increase in investment of 10 in Country 1 ($\Delta I_1 = 10$) will (*a*) raise income in Country 1 by 21, (*b*) raise income in Country 2 by 2.5, and (*c*) shift Country 1's balance of trade toward a deficit by 0.55. The full effects are shown in column 1 of Table 22–1.

TABLE 22–1

Foreign Trade Multipliers for Changes in Investment and Import Demand in a Two-Country Model

	(1) $\Delta I_1 = 10$, $\Delta \bar{M}_1 = 0$, $\Delta I_2 = 0$, $\Delta \bar{M}_2 = 0$	(2) $\Delta \bar{M}_2 = 10$, $\Delta I_1 = 0$, $\Delta I_2 = 0$, $\Delta \bar{M}_1 = 0$	(3) $\Delta I_2 = 10$, $\Delta I_1 = 0$, $\Delta \bar{M}_1 = 0$, $\Delta \bar{M}_2 = 0$	(4) $\Delta \bar{M}_1 = 10$, $\Delta I_1 = 0$, $\Delta I_2 = 0$, $\Delta \bar{M}_2 = 0$
$\Delta Y_1 (= \Delta C_1 + \Delta X_1 - \Delta M_1)$	21.00	11.00	10.00	−11.00
$\Delta C_1 (= 0.55\, \Delta Y_1)$	11.55	6.05	5.50	−6.05
ΔI_1	10.00	...	...	...
$\Delta X_1 (= \Delta M_2)$	0.50	5.50	5.00	4.50
$\Delta M_1 (= 0.05\, \Delta Y_1 + \Delta \bar{M}_1)$	1.05	0.55	0.50	9.45
$\Delta Y_2 (= \Delta C_2 + \Delta I_2 + \Delta X_2 - \Delta M_2)$	2.50	−22.50	25.00	22.50
$\Delta C_2 (= 0.78\, \Delta Y_2)$	1.95	−17.55	19.50	17.55
ΔI_2	...	...	10.00	...
$\Delta X_2 (= \Delta M_1)$	1.05	0.55	0.50	9.45
$\Delta M_2 (= 0.20\, \Delta Y_2 + \Delta \bar{M}_2)$	0.50	5.50	5.00	4.50
$\Delta B_1 (= \Delta X_1 - \Delta M_1)$	−0.55	4.95	4.50	−4.95
$\Delta B_2 (= \Delta X_2 - \Delta M_2)$	0.55	−4.95	−4.50	4.95
$\Delta Y (= \Delta Y_1 + \Delta Y_2)$	23.50	−11.50	35.00	11.50

Case 2: Multipliers Applicable to an Autonomous Change in the Demand for Imports in Country 2 **($\Delta \bar{M}_2$).** This is the same thing as an autonomous change in Country 1's exports. Its effects can be derived by setting $\Delta \bar{C}_1 = 0$, $\Delta I_1 = 0$, $\Delta \bar{M}_1 = 0$, $\Delta \bar{C}_2 = 0$, and $\Delta I_2 = 0$ in Equations 17 and 18 and solving the resulting equations for ΔY_1 and ΔY_2.[4] This yields

$$\Delta Y_1 = \frac{s_2}{s_1 s_2 + s_1 m_2 + m_1 s_2} \Delta \bar{M}_2 \qquad (23)$$

$$\Delta Y_2 = \frac{-s_1}{s_1 s_2 + s_1 m_2 + m_1 s_2} \Delta \bar{M}_2 \qquad (24)$$

The effect on Country 1's trade balance is given by

[4] It is assumed throughout this analysis that changes in import demand are at the expense of domestically produced consumption or investment goods. This is apparent in the above case from the fact that the increase in $\bar{M}_1$ is not accompanied by an increase in $\bar{C}_1$ or I_1. Since C and I include *all* spending on consumption and investment, whether the goods and services are produced at home or imported, an increase in imports with no increase in total consumption or investment implies a corresponding reduction in the demand for domestically produced goods and services.

$$\Delta B_1 = \Delta X_1 - \Delta M_1 = \Delta M_2 - \Delta M_1$$
$$\Delta B_1 = m_2\,\Delta Y_2 + \Delta\bar{M}_2 - m_1\,\Delta Y_1$$

Substituting the values of ΔY_1 and ΔY_2 from Equations 23 and 24 and simplifying we obtain

$$\Delta B_1 = \frac{s_1 s_2}{s_1 s_2 + s_1 m_2 + m_1 s_2}\,\Delta\bar{M}_2 \qquad (25)$$

For our numerical example, we have, upon substitution of $s_1 = 0.45$, $m_1 = 0.05$, $s_2 = 0.22$, and $m_2 = 0.20$ in Equations 23, 24, and 25:

$$\frac{\Delta Y_1}{\Delta\bar{M}_2} = 1.1$$
$$\frac{\Delta Y_2}{\Delta\bar{M}_2} = -2.25$$
$$\frac{\Delta B_1}{\Delta\bar{M}_2} = 0.495$$

That is, an autonomous increase of 10 in import demand in Country 2 (i.e., an autonomous increase in Country 1's exports) will (*a*) raise income in Country 1 by 11, (*b*) lower income in Country 2 by 22.5, and (*c*) shift Country 1's balance of trade toward a surplus by 4.95. The detailed effects of an increase of 10 in import demand in Country 2 are shown in column 2 of Table 22–1.

Case 3: Multipliers Applicable to an Autonomous Change in Investment in Country 2 (ΔI_2). Setting $\Delta\bar{C}_1 = 0$, $\Delta I_1 = 0$, $\Delta\bar{M}_1 = 0$, $\Delta\bar{M}_2 = 0$, and $\Delta\bar{C}_2 = 0$ in Equations 17 and 18 and solving the resulting equations for ΔY_1 and ΔY_2, we obtain

$$\Delta Y_1 = \frac{m_2}{s_1 s_2 + s_1 m_2 + m_1 s_2}\,\Delta I_2 \qquad (26)$$

$$\Delta Y_2 = \frac{s_1 + m_1}{s_1 s_2 + s_1 m_2 + m_1 s_2}\,\Delta I_2 \qquad (27)$$

The effect on the trade balance is given by

$$\Delta B_1 = \Delta X_1 - \Delta M_1 = \Delta M_2 - \Delta M_1$$
$$\Delta B_1 = m_2\,\Delta Y_2 - m_1\,\Delta Y_1$$

Substituting the values of ΔY_1 and ΔY_2 from Equations 26 and 27 and simplifying, we have

$$\Delta B_1 = \frac{s_1 m_2}{s_1 s_2 + s_1 m_2 + m_1 s_2}\,\Delta I_2 \qquad (28)$$

For our numerical example, we have, upon substituting the values, $s_1 = 0.45$, $m_1 = 0.05$, $s_2 = 0.22$ and $m_2 = 0.20$ in Equations 26, 27, and 28:

$$\frac{\Delta Y_1}{\Delta I_2} = 1$$

$$\frac{\Delta Y_2}{\Delta I_2} = 2.5$$

$$\frac{\Delta B_1}{\Delta I_2} = 0.45$$

Thus, an autonomous increase of 10 in investment demand in Country 2 would (*a*) raise income in Country 2 by 25, (*b*) raise income in Country 1 by 10, and (*c*) shift Country 1's trade balance toward a surplus by 4.5. Column 3 of Table 22–1 shows the full effects of an increase of 10 in investment demand in Country 2.

Case 4: Multipliers Applicable to an Autonomous Change in the Demand for Imports in Country 1 **($\Delta\bar{M}_1$).** This is the same thing as an autonomous change in Country 2's exports. Setting $\Delta\bar{C}_1 = 0$, $\Delta I_1 = 0$, $\Delta\bar{M}_2 = 0$, $\Delta\bar{C}_2 = 0$, and $\Delta I_2 = 0$ in Equations 17 and 18 we have

$$\Delta Y_1 = \frac{-s_2}{s_1 s_2 + s_1 m_2 + m_1 s_2} \Delta\bar{M}_1 \tag{29}$$

$$\Delta Y_2 = \frac{s_1}{s_1 s_2 + s_1 m_2 + m_1 s_2} \Delta\bar{M}_1 \tag{30}$$

The effect on Country 1's trade balance is given by

$$\Delta B_1 = \Delta X_1 - \Delta M_1 = \Delta M_2 - \Delta M_1$$
$$\Delta B_1 = m_2 \Delta Y_2 - m_1 \Delta Y_1 - \Delta\bar{M}_1$$

Substituting the values of ΔY_1 and ΔY_2 from Equations 29 and 30 and simplifying, we have

$$\Delta B_1 = \frac{-s_1 s_2}{s_1 s_2 + s_1 m_2 + m_1 s_2} \Delta\bar{M}_1 \tag{31}$$

For our numerical example, substituting the values $s_1 = 0.45$, $m_1 = 0.05$, $s_2 = 0.22$, and $m_2 = 0.20$ in Equations 29, 30, and 31, we have

$$\frac{\Delta Y_1}{\Delta\bar{M}_1} = -1.1$$

$$\frac{\Delta Y_2}{\Delta\bar{M}_1} = 2.25$$

$$\frac{\Delta B_1}{\Delta\bar{M}_1} = -0.495$$

Thus, an autonomous increase of 10 in import demand in Country 1 (i.e., an increase in Country 2's exports) will (*a*) lower income in Country 1 by 11, (*b*) raise income in Country 2 by 22.5, and (*c*) shift Country 1's trade balance toward a deficit by 4.95. The details are shown in column 4 of Table 22–1.

SOME IMPLICATIONS OF THE ANALYSIS

Several important conclusions concerning the relation between national income and international trade can be drawn from this analysis.

International Transmission of Instability

In an interdependent world such as that depicted by the model developed above, variations in the strength of investment demand in one country will not only cause fluctuations in income in that country but will spill over through induced variations in import demand to cause fluctuations in the rest of the world as well. For the country experiencing the initial change in investment, imports constitute a leakage from the domestic income stream, which has an effect similar to saving. Thus, a high marginal propensity to import in that country will reduce its domestic multiplier, thereby tending to make its economy more stable. At the same time, the effects of changes in investment are transmitted to the rest of the world through the induced change in imports. A high marginal propensity to import in the country experiencing the initial change in investment therefore means that a large impulse is transmitted to the rest of the world, thereby contributing substantially to instability there.

These matters can be clarified by reference to Table 22–2, which pulls together the multipliers for changes in investment in Country 1 and Country 2 and contrasts these with the multipliers that would prevail in the absence of foreign trade. (As was explained in Chapter 6, the multiplier for a change in investment in the absence of foreign trade is the reciprocal of the marginal propensity to save.) The upper section (A) of the table shows that in the absence of trade, the domestic multiplier for a change in investment in Country 1 would be 2.22. The presence of trade with the parameters assumed in our numerical example reduces this multiplier only slightly, to 2.1, and causes investment in Country 1 to have a rather small effect—a multiplier of 0.25—on income in Country 2. On the other hand, the lower section (B) of the table shows that the multiplier for a change in domestic investment in Country 2 in the absence of trade would be 4.55. The introduction of trade reduces this multiplier by 45 percent to 2.5, and causes the change in investment in Country 2 to have a substantial effect—a multiplier of 1—on income in Country 1. Thus, primarily because the marginal propensity to import in Country 2 (0.20) is much larger than that

TABLE 22–2
A. Multipliers for Change in Investment in Country 1

In the Absence of Foreign Trade		*With Foreign Trade*	
$\frac{\Delta Y_1}{\Delta I_1}$	$\frac{\Delta Y_2}{\Delta I_1}$	$\frac{\Delta Y_1}{\Delta I_1}$	$\frac{\Delta Y_2}{\Delta I_1}$
$\frac{1}{s_1}$	0	$\frac{s_2 + m_2}{s_1s_2 + s_1m_2 + m_1s_2}$	$\frac{m_1}{s_1s_2 + s_1m_2 + m_1s_2}$
When $s_1 = 0.45$		When $s_1 = 0.45$, $m_1 = 0.05$, $s_2 = 0.22$, $m_2 = 0.20$	
2.22	0	2.1	0.25

B. Multipliers for Change in Investment in Country 2

In the Absence of Foreign Trade		*With Foreign Trade*	
$\frac{\Delta Y_1}{\Delta I_2}$	$\frac{\Delta Y_2}{\Delta I_2}$	$\frac{\Delta Y_1}{\Delta I_2}$	$\frac{\Delta Y_2}{\Delta I_2}$
0	$\frac{1}{s_2}$	$\frac{m_2}{s_1s_2 + s_1m_2 + m_1s_2}$	$\frac{s_1 + m_1}{s_1s_2 + s_1m_2 + m_1s_2}$
When $s_2 = 0.22$		When $s_2 = 0.22$, $m_2 = 0.20$, $s_1 = 0.45$, $m_1 = 0.05$	
0	4.55	1.0	2.50

in Country 1 (0.05), the introduction of foreign trade reduces the domestic multiplier in Country 2 much more than it reduces the domestic multiplier in Country 1, and causes changes in investment in Country 2 to have much greater effects on income in the rest of the world than do changes in investment in Country 1.[5]

Relation between National Income and the Trade Balance

The relation between changes in a country's trade balance and changes in its national income is a rather complicated one which depends on the initiating cause. The following table shows the direction of the effects on income and the trade balance produced by autonomous changes in investment at home and abroad and by autonomous changes in exports and imports. The last column of the table indicates examples of each phenomenon in the numerical illustration given in Table 22–1. The term, "improve" as applied to trade balance means an increase in exports relative to imports and a movement in the direction of a surplus; the term "worsen" refers to a movement toward a deficit. Decreases in investment at home or abroad or

[5] This statement is somewhat oversimplified, since changes in the other parameters (the marginal propensities to save and to import in the rest of the world) as well as the domestic marginal propensity to import affect the extent to which the presence of foreign trade changes a country's domestic multiplier and also the extent to which fluctuations are transmitted to the rest of the world.

Initiating Cause	*Effect on Income*	*Effect on Trade Balance*	*Illustrations in Table 22–1*
Increase in investment demand at home	Increase	Worsen	Country 1 in column 1 Country 2 in column 3
Increase in investment demand abroad	Increase	Improve	Country 2 in column 1 Country 1 in column 3
Increase in home demand for imports	Decrease	Worsen	Country 1 in column 4 Country 2 in column 2
Increase in foreign demand for exports	Increase	Improve	Country 1 in column 2 Country 2 in column 4

decreases in exports or imports would, of course, produce changes in income and the trade balance opposite in direction to those indicated in the table.

An autonomous increase in domestic investment will increase income while moving the trade balance toward a deficit. In Table 22–1, column 1, an increase of 10 in investment demand in Country 1 raises that country's income by 21; this raises imports by 1.05, thereby causing income to rise in Country 2, which increases Country 2's imports (Country 1's exports) by 0.5. Thus, the increase of 21 in Country 1's income is associated with a movement of its trade balance toward a deficit by 0.55 ($\Delta B_1 = \Delta X_1 - \Delta M_1 = 0.50 - 1.05 = -0.55$). In all of the other cases—where the initiating cause is an autonomous change in investment abroad, in imports, or in exports—income and the trade balance move in the same direction; an increase in income is associated with an improvement in the trade balance, and a decrease in income is associated with a worsening of the trade balance. This is because, in these cases, the stimulus to income itself comes through the trade balance: an initial improvement or worsening of the trade balance is the cause of the rise or fall of income. Since, in practice, changes in investment at home and abroad and in outside forces affecting exports and imports are commonly all occurring simultaneously, it is not surprising that there is no easily predictable relation between changes in income and changes in the trade balance. Thus, for example, recession may occur in association with either a trade deficit or a trade surplus.

"Beggar-My-Neighbor" Policies

In terms of the analysis presented above, if a country is suffering from an inadequate level of demand and excessive unemployment, there are two ways in which it might proceed to correct the situation:

1. It might attempt to stimulate increased aggregate consumption and investment demand. This might be done by increasing government expenditures, by reducing taxes to stimulate consumption, or by the adoption of an easier monetary policy to stimulate investment spending.
2. It might attempt to stimulate its domestic economy by taking measures to increase exports or decrease imports. This might be done by depreciating its currency relative to the currencies of other countries, by subsidizing exports, or by raising trade barriers to discourage imports.

While policies of either type 1 or type 2 would be capable of increasing income and employment in the home country, their effects on the balance of payments and on the rest of the world would be quite different. Expansionary fiscal or monetary policies (type 1) would create new demand in the country adopting them; the resulting rise in income would increase imports, thereby shifting the balance of payments towards a deficit, transmitting a portion of the stimulus to the rest of the world, and raising incomes abroad. Thus, such policies would serve to raise total income for the world as a whole. This can be seen from Table 22–1. An increase of 10 in investment in Country 1 (column 1) raises income in Country 1 by 21 and in Country 2 by 2.5, thereby producing a total increase in world income of 23.5. An increase in investment of 10 in Country 2 (column 3) raises income in that country by 25 and raises income in Country 1 by 10 for a total increase in world income of 35.

On the other hand, measures that stimulate exports or discourage imports (type 2 above) will raise domestic income while shifting the balance of payments toward a surplus and depressing incomes abroad. In a sense the country adopting them does not add to total world demand, but rather stimulates its own economy by taking demand away from the rest of the world. This also can be seen from Table 22–1. If Country 1 stimulates its exports (the imports of Country 2) by 10 (colunm 2), this will raise the income of Country 1 by 11 while reducing the income of Country 2 by 22.5, thereby actually lowering income for the world as a whole by 11.5. If Country 2 stimulates its exports (the imports of Country 1) by 10 (column 4), the income of Country 2 will be raised by 22.5 while the income of Country 1 will be reduced by 11, with the result that the income of the world as a whole will be raised by 11.5.

It is apparent from these illustrations that when a country takes measures to stimulate exports or retard imports, these measures will definitely reduce incomes in other countries—unless, of course, they are offset by domestic action in those countries. Since these measures raise income at home and reduce it abroad, they may either raise or lower income for the world as a whole depending on which effect dominates. The change in income for the world as a whole (ΔY) is

$$\Delta Y = \Delta Y_1 + \Delta Y_2$$

To find the effect on world income of a change in Country 1's exports ($\Delta\bar{M}_2$), substitute the values of ΔY_1 and ΔY_2 given in Equations 23 and 24. This yields

$$\Delta Y = \frac{s_2 - s_1}{s_1 s_2 + s_1 m_2 + m_1 s_2} \Delta\bar{M}_2$$

ΔY will be positive if s_1 is less than s_2 and negative if s_1 is greater than s_2. That is, if the country that takes action to stimulate exports or retard imports has a smaller marginal propensity to save than the rest of the world, the increase in that country's income will exceed the decline in income in the rest of the world, and total world income will rise. On the other hand, if the country taking the action has a larger marginal propensity to save than the rest of the world, the increase in its income will be less than the fall in income in the rest of the world, and total world income will fall. That is, if a portion of aggregate demand is shifted from a country with a low multiplier (i.e., a high marginal propensity to save) to one with a high multiplier, total world income will be increased, and vice versa.

Certain important general principles of economic policy can be drawn from this analysis. It is generally not appropriate for a country to attempt to manipulate its foreign trade position as a means of stimulating economic activity at home. In the 1930's, many countries attempted to eliminate domestic unemployment by expanding exports or contracting imports through currency depreciation or other measures. When one country adopted measures of this kind, however, it merely served to accentuate the unemployment problems of other countries; when many countries tried to use such methods, the effects on income and employment tended to cancel out while the effects on world trade were chaotic. As a result of this experience in the 1930's, attempts to increase domestic employment by artificially stimulating exports or retarding imports have come to be called "beggar-my-neighbor" policies and are now generally regarded as improper. It is now recognized that the proper way to deal with domestic unemployment is through the application of expansionary fiscal and monetary policies rather than through manipulation of the trade balance.

ESTIMATION OF MULTIPLIERS

It should be noted that the presence of foreign trade complicates the estimation of the multiplier applicable to autonomous changes in domestic spending. In the case of the United States, fortunately, this complication is not too serious. It is customary to assume that the effects of changes in U.S. imports on the national income of other countries are offset by appropriate domestic monetary and fiscal policies in those countries. This is reasonable; it amounts to an assumption that other countries succeed in achieving the levels of aggregate demand they desire regardless of what we do. If this is the case, a change in investment in the United States does not change income in other countries and therefore does not change their imports from us. If we take the United States to be Country 1 in the model developed above, the effect of this assumption is to make m_2 effectively equal to zero; that is, since changes that occur in the United States do not affect income in other countries, they do not change their imports from us. If $m_2 = 0$, the effect of a change in domestic investment on income (Equation 20 above) reduces to

$$\Delta Y_1 = \frac{1}{s_1 + m_1} \Delta I_1$$

What this amounts to is that imports constitute an additional leakage from the domestic income stream, similar to saving (and taxes), which reduces the multiplier. In the case of the United States, however, the marginal propensity to import is normally only on the order of the 4 to 5 percent (i.e., $m_1 = 0.04$ to 0.05) so that the allowance for induced imports is of very minor importance in estimating the multiplier.[6]

In some other countries, such as the United Kingdom or the Netherlands, on the other hand, the marginal propensity to import is 30 to 40 percent, enough to affect the multiplier seriously. For a country such as Canada, which is tied closely to and to some extent dominated by the much larger economy of the United States, foreign economic conditions are often of predominant importance. Canada has to pay as much attention to developments in the United States as it does to those in its own country in deciding what policies to adopt.

EFFECTS OF PRICE CHANGES

The analysis developed above is oversimplified in its assumption that changes in aggregate demand cause changes in output and employment without any accompanying changes in prices. This assumption may be approximately correct when the economy is operating at low levels of employment with large quantities of unutilized resources. However, when the economy is in the neighborhood of full employment, the assumption is scarcely tenable. Under such circumstances, the above analysis gives an incomplete picture of the forces affecting the balance of payments.

An increase in aggregate demand in a country whose economy is close to full employment will cause an increase in prices as well as in production and employment. As indicated in connection with the analysis presented above,

[6] Actually, the average propensity to import of the United States (total imports divided by GNP) is typically 4 to 5 percent. In the long run, the average and marginal propensities to import are equal. In the short run, however, the marginal propensity varies considerably and may depart substantially from the average propensity.

the rise in real income associated with expanded production and employment will cause an increase in import demand to an extent measured by the marginal propensity to import. In addition, the rise in prices in the home country relative to those prevailing abroad will make home-produced goods less attractive relative to foreign-produced goods both in the domestic market and in foreign markets. This will cause an increase in imports and a decline in exports as both domestic and foreign buyers substitute the now relatively less expensive foreign-produced goods for home-produced goods.

Thus, the tendency for the rise in domestic income to move the balance of payments toward a deficit by increasing imports that is present in the model developed above will be reinforced by the effects of price changes, which will cause both a further increase in imports and a decline in exports. Moreover, the rise in imports (at the expense of home-produced goods) and the fall in exports brought about by the change in prices will offset some portion of the initiating autonomous rise in domestic expenditures, thereby weakening the effects on income and employment. At the same time, in the other country the increase in exports and the shift from imports to home-produced goods will add further to aggregate demand, thereby stimulating production and employment in that country. Further complications are introduced by the fact that if the other country is also close to full employment, the stimulus to aggregate demand will cause prices as well as income and employment to rise in that country also, and both the income and price changes will have favorable feedback effects on the home country. However, it is apparent that the rise in imports and the fall in exports caused by the increase in prices in the country initially experiencing the increase in aggregate demand will (*a*) worsen that country's balance of trade, (*b*) offset a portion of the increase in income and employment in the home country caused initially by the increase in aggregate demand, and (*c*) transmit part of the expansionary stimulus to the rest of the world, causing income and employment—and prices, if high employment prevails—to rise there also. The effects of price adjustments are thus closely parallel to those produced by changes in real income as described under *Case 1* in the previous section of this chapter and as depicted in column 1 of Table 22–1.

A formal two-country model describing the determination of income levels and the balance of trade taking account of both income and price changes would be extremely complex, and we shall not attempt to develop one. Such a model is not really necessary for our purposes in any case. In general, changes in real income are accompanied by changes in prices in the same direction. Moreover, price changes affect the balance of trade in a manner similar to income changes: a rise in income worsens the balance of trade by stimulating imports, while a rise in domestic prices (relative to foreign prices) worsens the balance of trade by both stimulating imports and retarding exports. Accordingly, as the above discussions of the effects and of an autonomous increase in aggregate demand illustrates, changes in prices merely serve to accentuate the effects on the balance of trade, on domestic income, and on foreign income that are present in a model such as that developed in the previous section, in which price changes are not taken into account. This being the case, that model will serve our purposes quite satisfactorily.

Chapter 23 METHODS OF DEALING WITH BALANCE-OF-PAYMENTS DISEQUILIBRIA

We may describe a country's balance of payments as being in equilibrium if it has neither a deficit nor a surplus calculated on an official reserve transactions basis—that is, if its international reserve position is neither improving nor worsening. In this chapter we shall first discuss briefly some of the causes of balance-of-payments disquilibria and then the methods that may be used to deal with such disequilibria once they have arisen.

TYPES AND SOURCES OF BALANCE-OF-PAYMENTS DISEQUILIBRIA

There are many forces that might cause a country's balance-of-payments position to depart from equilibrium in the direction of either a deficit or a surplus. While it would be impossible to catalog all of the possible causes of deficits or surpluses, it will be useful to classify the sources of balance-of-payments disturbance under a few headings and to refer to a few of the specific causes that fall under each heading.

Temporary Disequilibria

Obviously, even if a country's balance of payments is fundamentally in balance, inpayments and outpayments will not match on a day-to-day basis. Outpayments may exceed inpayments for a short time, followed by a period in which the opposite situation prevails. Temporary disequilibria of this kind may be caused by random variations in trade, seasonal fluctuations, the effects of weather on agricultural output, and so on. Deficits or surpluses resulting from such temporary causes can ordinarily be expected to reverse themselves within a fairly short time.

Fundamental or Chronic Disequilibria

Fundamental or chronic disequilibria are due to basic changes in the economic position of a country. They may result from shifts in consumer tastes at home or abroad which affect the country's imports or exports; from technological improvements in products or methods of production in the country's industries, or in the industries of other countries, which affect the country's ability to compete either in the home market or in foreign markets; or from differences in the rates of growth of the labor force or of capital accumulation between the country and its competitors.

Sudden major changes in a country's international economic and financial position have often been a source of serious problems of balance-of-payments adjustment. Frequently, these sudden structural changes have been a result of major wars. Wartime destruction of productive capacity may reduce a country's ability to export and also its capacity to meet its domestic needs from its own production, thus requiring increased imports. Examples include the bomb damage and destruction of British merchant ships during World War II, which drastically reduced British foreign exchange earnings, and the heavy destruction of productive capacity in Germany also during World War II. War may also lead to heavy liquidation of foreign investments as a means of obtaining foreign exchange to pay for the import of war materials and also to the piling up of heavy international indebtedness as a means of financing the war effort.

Examples include the heavy British liquidation of holdings of American securities, the building up of debts by Britain to other members of the Sterling Area during World War II, and the huge accumulation of war debts to the United States by other allied countries during World War I. Developments of this kind lead to a drastic increase in outpayments on capital account and for the payment of interest. The maintenance of balance-of-payments equilibrium requires that such large increases in outpayments be covered by increased earnings from other sources, such as exports. War may also lead to the imposition of heavy reparations payments on the defeated countries by those who are victorious. One of the major sources of balance-of-payments difficulties in the 1920's was the heavy reparations burden imposed on Germany after World War I. Finally, war may lead to an extensive destruction of markets and disruption of normal economic relationships. For example, the destruction of the German economy in World War II left a major vacuum in world trade, requiring the establishment of new trade relationships.

Foreign policy considerations may at times impose heavy strains on a nation's balance of payments which call for important economic readjustments. One of the causes of the chronic deficits that have characterized the U.S. balance of payments in recent years has been heavy outlays for foreign aid grants and loans and military expenditures overseas. If such expenditures are undertaken and maintained for a considerable period of time, they may be a source of fundamental or chronic disequilibrium in a nation's balance of payments.

Cyclical Changes in the Balance of Trade

Deficits and surpluses in the balance of trade may occur from time to time as a by-product of cyclical fluctuations in economic activity. However, the way in which a country's trade balance will behave depends upon whether the fluctuations originate at home or abroad. For example, if a recession occurs as a result of a decline in investment spending in the United States, our income will fall, causing imports to decline, thereby moving the U.S. trade balance toward a surplus—and transmitting the recession to other countries. If, on the other hand, the recession is initiated through a decline in investment spending abroad, the fall in income abroad will cause U.S. exports to decline, thereby shifting the U.S. trade balance toward a deficit in the process of transmitting the recession to the United States. Thus, cyclical changes in business activity are accompanied by characteristic changes in the balance of trade, which may be classified as follows:

Cyclical Change	*Accompanying Change in the Balance of Trade*
Depression starting at home	Improvement
Inflation starting at home	Deterioration
Depression starting abroad	Deterioration
Inflation starting abroad	Improvement

If changes originate simultaneously at home and abroad, the above classification relates to whether the forces at work are stronger at home or abroad. Cyclical changes are of a nature intermediate between temporary disequilibria and chronic or fundamental disequilibria. They may pose quite complex problems of balance-of-payments policy.

Final Comment

While the above classification is useful for analytical purposes, it should be noted that the balance of payments is *simultaneously buffeted by many forces*—some temporary, some fundamental, some cyclical in nature. It is often impossible to distinguish clearly among them, due to the complexity and simultaneity of economic forces.

BALANCE-OF-PAYMENTS POLICIES

A country which experiences a balance-of-payments deficit or surplus has a choice of a number of ways of dealing with the situation. In the discussion of these alternatives below, primary attention is paid to methods of handling deficits, since (for reasons that will be explained shortly) deficits commonly impose greater strains on countries experiencing them than do surpluses.

Financing the Deficit

A country may permit a deficit to continue without taking measures to reduce it and simply draw on its gold and foreign exchange reserves to finance the deficit while it waits for the operation of "natural economic forces" to reverse its position. This is generally the best way to deal with temporary deficits resulting from random variations in trade, seasonal fluctuations, or crop failures. It may also be a sensible way to handle a deficit that is clearly attributable to a cyclical change in business activity that is likely to be reversed within a relatively short time. However, the extent to which this method can be used clearly depends upon (*a*) the size of the country's reserves, and (*b*) the magnitude and duration of the disturbance. Moreover, if reserves are to be drawn upon systematically to deal with reversible deficits, the country must use the surpluses that occur during ensuing periods to rebuild its reserves.

Under an international monetary system of the kind that now exists, there is an important asymmetry between deficits and surpluses. Since monetary reserves cannot become negative, it is clear that a country can only continue to finance deficits until such time as its monetary reserves have been used up —subject, of course, to the qualification that it may, to a limited extent, be able to borrow additional reserves from the IMF or from other countries. There is, however, no equivalent limitation on the ability of the country to "finance" a balance-of-payments surplus by accumulating reserves. This asymmetry is important because it means that a country experiencing a deficit is under greater pressure to take some kind of corrective action than is a country experiencing a surplus.

The above discussion of the limits on the ability of the country to finance a deficit applies to a nonreserve currency country. For a country such as the United States, whose currency forms part of the reserves of other countries, the situation is a little more complicated and the possibilities of financing a deficit may be considerably greater. A reserve currency country will be able to finance a deficit without drawing down its holdings of gold and foreign exchange if foreign central banks and governments are willing to increase their holdings of its currency. For example, the United States experienced large deficits during the 1950's which it had no difficulty in financing because other countries were anxious to rebuild their monetary reserves, which had been depleted in the course of World War II, and they were therefore willing to absorb large quantities of dollars. By the beginning of the 1960's, however, foreign governments had succeeded in rebuilding their reserves to reasonably adequate levels. At the same time foreign official holdings of short-term dollar claims had risen substantially relative to the U.S. gold stock available for redeeming them, and, as a consequence, foreign central banks and governments became less willing to accumulate additional dollars. Nevertheless, it has continued to be possible in the last few years for the United States to finance some portion of its continuing large deficits by the emission of dollars which have been absorbed by foreign central banks and governments. Indeed, our difficulties in financing the large deficits of the 1960's would unquestionably have been substantially greater than they have been had the dollar not been a reserve currency.

Restoration of Equilibrium through Changes in Incomes, Prices, and Interest Rates

It would generally be possible, at least in principle, to eliminate a balance-of-payments deficit or surplus by means of appropriate internal changes in incomes, prices, and interest rates. These changes might be brought about by the adoption of suitable fiscal and monetary policies. For example, a nation having a balance-of-payments deficit might attempt to eliminate the deficit by adopting restrictive fiscal and monetary policies—that is, by raising taxes, reducing government expenditures, reducing the stock of money to raise interest rates, or by adopting some combination of these measures. An increase in taxes or a reduction in government expenditures would reduce aggregate demand, income, production, and employment, and, to the extent that import demand was determined by income, would reduce imports. The decline in imports would impart deflationary pressures to other countries, which, unless offset by domestic policies in those countries, would lead to a reduction in the initiating country's exports. However, as our earlier discussion showed, the decline in imports would exceed that in exports, and the balance-of-payments deficit would therefore be reduced. A restrictive monetary policy would have a similar effect on the balance of trade, as higher interest rates reduced domestic investment, income, employment, and exports. In addition, to the extent that a restrictive monetary policy raised interest rates in the home country relative to those prevailing in other countries, capital might be induced to flow into the country in search of a high return. Thus, a tightening of monetary policy might improve both the trade account and the capital account of the balance of payments. Restriction of aggregate demand brought about by fiscal or monetary policies might also reduce prices at home relative to those prevailing abroad, and such price changes would complement the effects of income changes in improving the balance of trade. Conversely, a surplus in the balance of payments might be corrected by the adoption of policies opposite to those called for to correct a deficit. An expansionary fiscal policy, implemented by cutting taxes or increasing government expenditures, would increase income and raise prices, thereby reducing the surplus in the trade balance by increasing imports and reducing exports. An expansionary monetary policy would have a similar effect on the balance of trade, and, in addition, by lowering interest rates at home compared with those abroad, might cause capital to flow out of the country.

Automatic Corrective Forces

Actually, the appearance of a deficit or surplus in the balance of payments will *automatically* set in motion forces working toward the restoration of equilibrium. Suppose, for example, that a country experiences an autonomous decline in foreign demand for its exports, caused, let us say, by a change in tastes on the part of foreign buyers. As was shown in our earlier exposition of the foreign trade multiplier, such a decline in exports—unless offset by other measures—would (*a*) reduce income, production, and employment in the exporting country, and (*b*) increase income, production, and employment in other countries. These changes in income at home and abroad would, in turn, bring about changes in both exports and imports that would cancel out a portion of the effect on the balance of payments produced by the initial drop in exports. This can be seen by reference to our earlier numerical illustration of the working of the foreign-trade multiplier. The effect of the autonomous decline of 10 in the demand for Country 1's exports can be seen by reversing the signs of the entries in column 2 of Table 22–1. In this case, the decline in Country 1's exports causes income in Country 2 to rise by 22.5 (because it is assumed that the fall in Country 2's demand for imports from Country 1 is associated with a shift in demand toward home-produced goods in Country 2). Since Country 2's marginal propensity to import is 20 percent, the rise in income causes an increase of 4.5 (0.2×22.5) in Country 2's imports (Country 1's exports). Thus, Country 1's exports fall by only 5.5 (the original drop of 10 minus the induced increase of 4.5). Moreover, the fall in Country 1's exports causes income in that country to decline by 11, and, with a marginal propensity to import of 5 percent, the imports of Country 1 fall by 0.55 (0.05×11). Thus, the autonomous decline of 10 in Country 1's exports causes that country's balance of payments to shift toward a deficit to the extent of only 4.95 (the initial decline in exports of 10 minus the induced increase in exports of 4.5 minus the induced fall in imports of 0.55). Although the precise results depend, of course, on the marginal propensities to save and to import in the two countries, the income effects working through the foreign-trade multiplier, provided policies are not adopted to offset them, will produce a partial correction of the deficit or surplus caused by an autonomous change in export or import demand. Price effects associated with the expansion or contraction of income in the countries involved may make some further contribution to the correction; in the above illustration, there would probably be some fall in prices in Country 1 and some increase in prices in Country 2, which would further reduce Country 1's imports and increase its exports.

In addition to the direct *income effects* referred to above, the appearance of a deficit or surplus produces *monetary effects* in the two countries, which, unless offset by other policy actions, will contribute a further automatic corrective force. A country experiencing a balance-of-payments deficit will find that the demand for foreign exchange to purchase goods, services, and securities exceeds the supply of foreign exchange generated by sales of goods and services and securities, with the result that the foreign exchange value of its currency will fall. Thus, it will soon find itself in the position of England as depicted in Panel C of Figure 21–2, and its central bank will have to sell foreign exchange from the country's reserves in order to prevent the exchange rate from falling more than 1 percent below the IMF parity. These sales of foreign exchange by the central bank will reduce the reserves of the commercial banking system in the same way as would an open-market sale of government securities. The reduction in bank reserves will force a contraction of the supply of money and bank credit, thereby driving up interest rates and reducing investment demand. On the other hand, a country experiencing a balance-of-payments surplus will find the supply of foreign exchange resulting from exports of goods, services, and securities in excess of the demand for foreign exchange to pay for imports of goods, services, and securities. This will cause the foreign exchange value of its currency to rise, putting it in the position depicted for England in Panel B of Figure 21–2. In order to prevent the exchange rate from rising more than 1 percent above the IMF parity, the surplus country's central bank will have to enter the market and buy foreign exchange. These purchases of foreign exchange by the central bank will increase the supply of bank reserves in the same manner as would an open-market purchase of government securities. The increase in bank reserves will provide the basis for an expansion of the money supply and a decline in interest rates, thereby leading to an increase in investment demand and an expansion of income, employment, and prices; and this, in turn, will bring an increase in imports and a reduction in exports by the surplus country, thereby moving its balance of payments back toward equilibrium. Moreover, the rise in interest rates in the deficit country combined with the fall in interest rates in the surplus country may induce a flow of interest-sensitive capital from the surplus country to the deficit country. Thus, these monetary adjustments will affect both the trade account and the capital account in the direction of restoring equilibrium. And, since the adjustments will persist as long as the countries involved continue to experience deficits and surpluses, they are capable, in principle, of producing a full restoration of equilibrium in the course of time provided they are allowed to operate without interference. Automatic monetary adjustments produced by gold movements from deficit countries to surplus countries were, of course, the adjustment mechanism that was emphasized under the classical gold standard. It is interesting to note that these forces are still potentially present under the international monetary system now in operation.

Offsetting versus Reinforcing Actions

It is possible for a country to *offset* the automatic corrective forces that are set in motion by a balance-of-payments deficit or surplus. A reduction in export demand can be offset by an increase in government expenditures or by a reduction in tax rates or interest rates to increase private domestic consumption or investment demand. If a deficit is causing a reduction in the domestic money supply, this can be offset by expanded purchases of securities in the open market by the central bank. Similarly, the automatic expansionary effects of a surplus can be offset by the adoption of a more restrictive fiscal and monetary policy—by reducing government expenditures, raising tax rates, or selling government securities in the open market. Alternatively, instead of offsetting the automatic corrective forces, a country may follow a policy of *neutrality,* permitting them to operate without interference. Or it may even *reinforce* them: for example, a country having a deficit may strengthen the corrective forces by reducing government expenditures, raising tax rates, or selling government securities in the open market.

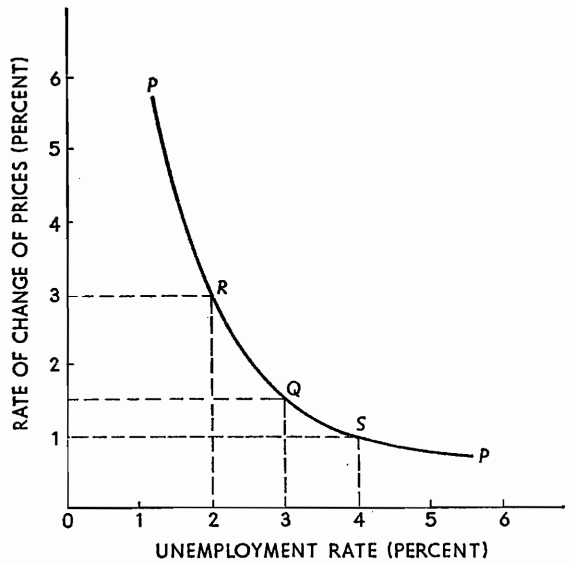
External versus Internal Stability

Under some circumstances the corrective action that is appropriate for the balance of payments is also right for the domestic economy, but this is by no means always the case. Suppose the modified Phillips curve of a country is as depicted by *PP* in Figure 23–1, and suppose the point selected as optimal by the authorities responsible for economic policy is point *Q*.[1] The country may actually be at a point such as *R,* at which it is experiencing excessive inflation—a situation sometimes described as "overfull employment"—or it may be at a point such as *S,* at which it is experiencing excessive unemployment. Each of these situations may occur in combination with either a deficit or a surplus in the balance of payments. The resulting four cases are shown in Table 23–1, the last two columns of which indicate the kind of fiscal and monetary policy called for by the balance of payments and by the domestic economic situation. In Case I, in which the country has a deficit and is at a point (such as *R* of Figure 23–1) where it is experiencing excessive inflation, both the domestic economy and the balance of payments call for a restrictive fiscal and monetary policy. Such a policy will improve the trade

[1] The modified Phillips curve is discussed in Chapter 16, pp. 366–68, and in Chapter 17, pp. 381–83.

FIGURE 23–1

Modified Phillips Curve



balance by lowering real income and by slowing the increase in prices, thereby leading to reduced imports and increased exports; if a restrictive monetary policy is employed, it may also improve the balance on capital account as higher interest rates attract an inflow of funds from abroad. On the other hand, in Case IV, which involves a balance-of-payments surplus combined with a domestic situation of excessive unemployment such as

TABLE 23–1

Fiscal and Monetary Policies for Internal and External Stability

Case	*Balance of Payments Situation*	*Domestic Economic Situation*	*Fiscal and Monetary Policy Appropriate for* — *Balance of Payments*	*Domestic Economy*
I	Deficit	Excessive inflation (point *R*, Figure 23–1)	Contractionary	Contractionary
II	Deficit	Excessive unemployment (point *S*, Figure 23–1)	Contractionary	Expansionary
III	Surplus	Excessive inflation (point *R*, Figure 23–1)	Expansionary	Contractionary
IV	Surplus	Excessive unemployment (point *S*, Figure 23–1)	Expansionary	Expansionary

depicted by point S in Figure 23–1, an expansionary fiscal and monetary policy is called for in the interest of both the domestic economy and the balance of payments.

In Cases II and III, however, there are conflicts between the requirements of *internal stability* (an acceptable combination of unemployment and inflation) and the requirements of *external stability* (balance-of-payments equilibrium).[2] In Case II (balance-of-payments deficit combined with excessive unemployment), the balance of payments calls for a restrictive fiscal and monetary policy while the domestic economic situation calls for an expansionary policy. In Case III (surplus combined with excessive inflation), the situation is just the reverse: the balance of payments calls for an expansionary policy while the domestic economy calls for a restrictive policy. From a practical standpoint, situations of conflict of this kind are not uncommon. Most of the time during the first half of the 1960's, the United States was in the Case II situation, having excessive unemployment and a deficit, while the countries of continental Western Europe, notably Germany and France, were often in the Case III situation, having excessive inflation combined with a surplus.

Use of Trade or Exchange Controls

Measures such as import duties, import quotas, and export subsidies may be used to deal with balance of payments. Such measures of "commercial policy" often have other objectives: encouraging development, protecting living standards (largely fallacious), protecting favored domestic industries, etc. But they may also be used—and have been used in many instances—to protect or adjust the balance of payments. Alternatively, exchange controls have often been used for balance-of-payments reasons. These usually involve mandatory sale of foreign exchange proceeds derived from exports to a government authority and purchase of exchange from the authority under a system of licensing to pay for imports. The authorities may use nonprice rationing of foreign exchange or discriminatory pricing practices to discourage imports and encourage exports when the country is threatened with a balance-of-payments deficit. The controls may apply to residents of the country, to nonresident foreigners, or to both. Exchange controls may be used to regulate capital movements as well as trade in goods and services.

Exchange Rate Adjustments

Another way to eliminate a balance-of-payments deficit or surplus is to make an appropriate adjustment in the exchange rate. Under the IMF charter a member country is permitted to make an adjustment in its ex-

[2] For a detailed discussion of conflicts between internal and external stability, see James E. Meade, *The Balance of Payments* (London: Oxford University Press, 1951), chap. X.

change parity to deal with a "fundamental disequilibrium" in its balance of payments. Such an adjustment ordinarily requires the approval of the IMF.

Effects of Exchange Depreciation. Depreciation of a country's currency—that is, a reduction in its par value in terms of other currencies—will ordinarily improve the country's balance of trade by reducing foreign exchange payments for imports and increasing foreign exchange receipts from exports. As indicated earlier in this chapter, the U.S. dollar is now defined as 0.888671 grams of gold, which determines the official price of gold at $35 an ounce. Suppose the United States were to raise the official price of gold by 25 percent, to $43.75 an ounce.[3] If the gold content of the British pound remained unchanged at 2.13281 grams, the exchange parity between the dollar and sterling would rise by 25 percent from $2.40 to $3.[4] Or, to put it another way, the dollar would be depreciated by 20 percent in terms of sterling (and other currencies whose parities remained unchanged), from $\$1 = £5/12$ to $\$1 = £1/3$. To illustrate the effects of this change, consider a British commodity having a domestic price of £1. Before the depreciation, this commodity sold for $2.40 in the United States. After the depreciation, assuming its internal price remains at £1, its price in the United States will rise by 25 percent to $3. On the other hand, consider a U.S. export whose internal price is $2.40. Before the depreciation of the dollar, this commodity sold for £1 in Britain; after the depreciation, assuming its internal price remains unchanged, its price in Britain will fall by 20 percent to £0.8. Thus, the depreciation raises the prices of British goods in the United States by 25 percent and lowers the prices of U.S. goods in Britain by 20 percent. The effect of these changes would be to retard U.S. imports and stimulate U.S. exports.

The effect of the depreciation on the U.S. balance of payments might be as depicted in Figure 23–2. At the original exchange parity of $2.40, U.S. foreign exchange receipts from exports are *OA,* while payments of foreign exchange for imports are *OB;* this leaves a deficit of *AB.* By lowering the prices of U.S. goods in Britain, a 25 percent increase in the price of gold which raises the exchange parity to $3 increases the foreign exchange proceeds from U.S. exports by *AA';* by raising the prices of British goods in the United States, the depreciation reduces the U.S. demand for sterling by *BB'.* Thus, the deficit is reduced from *AB* to *A'B'.* It may be noted that even after the depreciation, the exchange rate is below equilibrium. The equilibrium exchange rate—the rate at which supply and demand are equated and the deficit is entirely eliminated—is $3.60 in Figure 23–2.

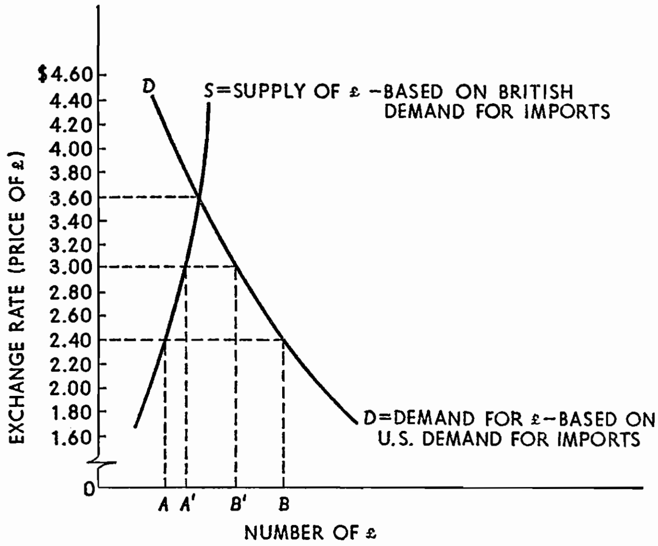
A Possible Complication. Although the situation in which depreciation reduces the depreciating country's payments of foreign exchange and increases its receipts, as depicted in Figure 23–2; is probably the normal case,

[3] Since there are 31.103481 grams in an ounce, this would involve a reduction of 20 percent in the gold content of the dollar, from 0.888671 grams ($31.103481 \div 35 = 0.888671$) to 0.710937 ($31.103481 \div 43.75 = 0.710937$).

[4] That is, $2.13281 \div 0.888671 = 2.4$, while $2.13281 \div 0.710937 = 3$.

FIGURE 23–2

Effects of Exchange Depreciation



it is not quite certain that things will work out this way. Indeed, it is not completely impossible that depreciation will make the depreciating country's deficit larger rather than smaller. To see how this might happen, we will need to examine the demand side and the supply side of the foreign exchange market separately.

The demand for foreign exchange can be dismissed rather quickly. Since depreciation makes foreign exchange more expensive, we can be sure that it will reduce the quantity demanded by the importing country. That is, we can be sure that the demand curve for pounds (*DD*) in Figure 23–2 will have a negative slope.

There is, however, more uncertainty about the effect of depreciation on the supply of foreign exchange derived from the depreciating country's exports. While the depreciating country can be sure that it will sell a larger quantity of *goods* as a result of the depreciation, it is possible that it may receive *less* foreign exchange. This can perhaps best be brought out by a numerical example. Suppose the situation with respect to U.S. exports of a commodity before and after the depreciation is as follows:

	(1) *Exchange Rate*	(2) *U.S. Export Price*	(3) *British Price* (2) ÷ (1)	(4) *Quantity Demanded in Britain*	(5) *Total U.S. Receipts of £* (3) × (4)
Before depreciation	$2.40	$24	£10	1,000	£10,000
After depreciation	3.00	24	8	1,100	8,800

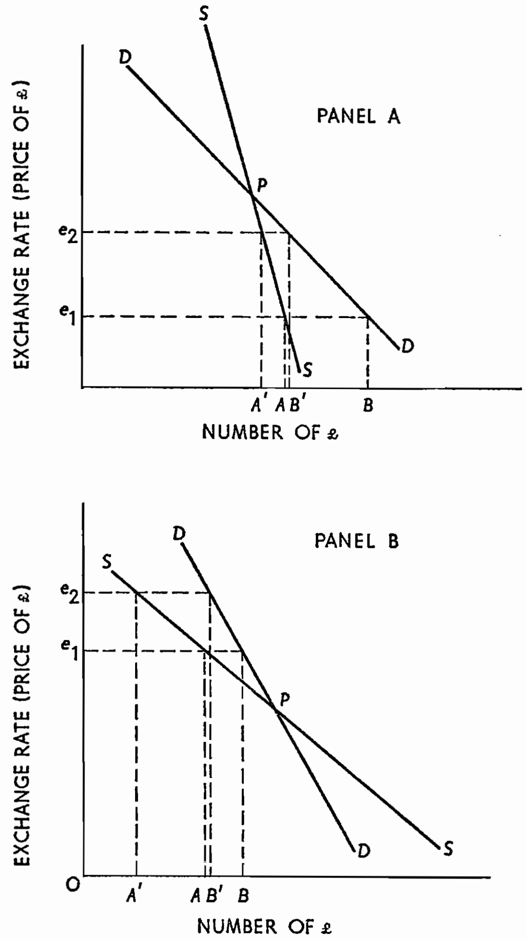
In this case, the depreciation lowers the price of the commodity in Britain from £10 to £8, and this causes British users to increase their purchases of the commodity from 1,000 units to 1,100 units. However, they spend fewer pounds on the good—£8,000 as compared with £10,000—after the depreciation than they did before. This result is attributable to the fact that the British demand for this particular good is inelastic—that is, of less than unitary elasticity. Thus, if the demands for the various categories of goods and services exported by the depreciating country are typically inelastic, exchange depreciation may reduce the depreciating country's export proceeds.

The supply of foreign exchange (curve *SS* in Figure 23–2) is based on the foreign demand for the depreciating country's export goods. As indicated above, if foreign demand for the country's exports is inelastic, depreciation reduces the foreign exchange proceeds derived from exports. It follows that in this case the supply of foreign exchange will have a negative slope instead of the positive slope exhibited in Figure 23–2. Cases in which the supply of foreign exchange has a negative slope—or in which, as it is sometimes stated, the supply curve is "backward rising"—are depicted in Figure 23–3. There are two such cases, as indicated in Panel A and Panel B of that figure. In Panel A the supply curve has a negative slope that is steeper than that of the demand curve. In this case, the negative slope of the supply curve does not prevent depreciation of the dollar (depicted by an increase in the exchange rate) from improving the U.S. balance of payments. When the exchange rate is e_1 in Panel A, the deficit is *AB.* If the dollar is depreciated to e_2, the proceeds from exports decline by *AA′,* while the foreign exchange payments for imports decrease by *BB′*. Since the decline in payments (*BB′*) exceeds the decline in receipts (*AA′*), the U.S. deficit is reduced from *AB* to *A′B′*.

In Panel B, on the other hand, the demand for foreign exchange has a steeper (negative) slope than the supply. In this case, depreciation will actually worsen the U.S. balance of payments. At exchange rate e_1, the deficit is *AB.* Depreciation of the dollar to e_2 reduces export proceeds by *AA′* and reduces payments for imports by *BB′*. Since the reduction in export proceeds is greater than the reduction in import payments, depreciation of the dollar from e_1 to e_2 increases the U.S. deficit from *AB* to *A′B′*. It may be noted that in this case deficits occur at exchange rates above the equilibrium level (point *P*) whereas in the other cases (Figure 23–2 and Panel A of Figure 23–3) deficits occur at rates below equilibrium. Thus, in the case depicted in Panel B of Figure 23–3, if the United States is experiencing a deficit, exchange *appreciation* (a reduction in the exchange rate) will be required to reduce or eliminate it. While situations of this kind may conceivably be encountered on occasion, they should probably be regarded as somewhat pathological; normally depreciation can be expected to improve the balance of payments, as in Figure 23–2 or Panel A of Figure 23–3.

FIGURE 23–3

Cases of Negatively Sloped Supply Curve of Foreign Exchange



Changes in Internal Prices. The analysis developed above is oversimplified in its assumption that the home prices of the depreciating country's exports and the foreign prices of its imports do not change as a result of depreciation. In fact, since the depreciation initially increases the foreign demand for the country's exports, home prices of exports are likely to rise, the extent of the increase depending on supply conditions in the export industries. Similarly, since the depreciation reduces the demand for imports, the prices of these goods in the countries in which they are produced are likely to decline, the amount of the decline depending on the supply conditions in the export industries of these countries. The analysis presented above applies to the special case in which supplies are infinitely elastic in export industries both at home and abroad so that internal prices do not change.

Actually, it can be shown that, in the absence of income effects (discussed below), exchange depreciation will improve the balance of payments provided

$$\eta_x + \eta_m > 1$$

where η_x is the price elasticity of foreign demand for exports (with sign reversed)[5] and η_m is the price elasticity of home demand for imports (with sign reversed). This condition holds in slightly attenuated form even in cases in which the supplies of export goods are not completely price-elastic at home and abroad.[6] Thus, as indicated above, if the foreign demand for exports is inelastic so that the supply curve of foreign exchange is "backward rising," this need not mean that exchange depreciation will worsen the balance of payments provided the home demand for imports has a sufficient degree of elasticity. If, for example, the elasticity of foreign demand for exports is -0.2 ($\eta_x = 0.2$) and the elasticity of home demand for imports is -0.9 ($\eta_m = 0.9$), we have

$$\eta_x + \eta_m = 0.2 + 0.9 = 1.1$$

so that depreciation will improve the balance of payments.

Effects on Income and Employment. Another oversimplification in the analysis presented above lies in its assumption that income and employment are constant and unaffected by an adjustment in the exchange rate. The

---

[5] In the numerical example on page 484, the elasticity of foreign demand for the U.S. export may be calculated by the usual arc elasticity formula, as follows (where q_0 is the quantity purchased before the price change, q_1 is the quantity purchased after the price change, p_0 is the initial price, and p_1 is the price after the change):

$$\text{Elasticity of demand} = \frac{\dfrac{q_1 - q_0}{q_1 + q_0}}{\dfrac{p_1 - p_0}{p_1 + p_0}} = \frac{\dfrac{1{,}100 - 1{,}000}{1{,}100 + 1{,}000}}{\dfrac{8 - 10}{8 + 10}} = -0.43$$

Thus, in this instance, after the reversal of the sign, we have $\eta_x = 0.43$.

[6] When supplies are completely elastic, $\eta_x + \eta_m < 1$ is both a necessary and a sufficient condition for exchange depreciation to worsen the balance of payments. On the other hand, when supplies are not completely elastic, it is a necessary but not a sufficient condition—that is, depreciation cannot worsen the balance of payments if $\eta_x + \eta_m > 1$, but even if $\eta_x + \eta_m < 1$, depreciation may improve the balance of payments provided supply conditions are such as to overcome the effects of the low demand elasticities. These generalizations hold true only when trade was balanced initially (i.e., before the currency was depreciated); if there was a deficit or a surplus initially, the situation is somewhat more complicated. See L. A. Metzler, "The Theory of International Trade," in H. S. Ellis (ed.), *A Survey of Contemporary Economics* (Philadelphia: Blakiston Co., 1948), Vol. I, pp. 225–28; W. L. Smith, "Effects of Exchange Rate Adjustments on the Standard of Living," *American Economic Review*, Vol. XLIV, December 1954, pp. 808–25.

fact is that an exchange rate adjustment may have important effects on income and employment both in the country making the adjustment and in other countries. Although extraordinarily complex analysis is required to take full account simultaneously of both the price and income effects, it is possible to indicate in a general way the effects of a variable level of employment on the results of an exchange rate adjustment.[7] Consider an exchange depreciation under conditions in which the supplies of export goods are perfectly elastic in both countries and suppose that the condition explained above

$$\eta_x + \eta_m > 1$$

is satisfied, so that the trade balance of the depreciating country is improved (and that of the rest of the world is, of course, worsened). This improvement in the depreciating country's trade affects the levels of income in that country and in the rest of the world in accordance with the multipliers for autonomous changes in exports and imports given in Chapter 22 (Equations 23 and 24). These multipliers are (with subscript 1 indicating the depreciating country)

$$\Delta Y_1 = \frac{s_2}{s_1 s_2 + s_1 m_2 + m_1 s_2} \Delta \bar{B}_1$$

$$\Delta Y_2 = \frac{-s_1}{s_1 s_2 + s_1 m_2 + m_1 s_2} \Delta \bar{B}_1$$

Here $\Delta \bar{B}_1$ is the improvement in the balance of payments that is initially caused by the depreciation (the decrease in imports plus the increase in exports). The logic behind this result is easy to see: In the depreciating country aggregate demand is increased both by the rise in exports and by the shift in demand from imports to domestically produced goods that is caused by depreciation. Conversely, in the rest of the world, aggregate demand is depressed both by the fall in exports and by the shift in demand from domestically produced goods to imports. These initial changes in aggregate demand set off multiplier effects which increase income and employment in the depreciating country and reduce income and employment in the rest of the world. It may be noted that the income effects will cancel out a portion of the improvement in the balance of payments that would have occurred in their absence. This can be seen from the expression (see Chapter 22, Equation 25):

$$\Delta B_1 = \frac{s_1 s_2}{s_1 s_2 + s_1 m_2 + m_1 s_2} \Delta \bar{B}_1$$

Here $\Delta \bar{B}_1$ is the improvement in the balance of payments produced by the exchange depreciation in the absence of changes in income, while ΔB_1 is

[7] A formal analysis of this problem is to be found in R. W. Jones, "Depreciation and the Dampening Effect of Income Changes," *Review of Economics and Statistics,* Vol. XLII, February 1960, pp. 74–80.

the ultimate change in the balance of payments after allowing for income effects. Since $s_1s_2/(s_1s_2 + s_1m_2 + m_1s_2) < 1$, ΔB_1 is smaller than $\Delta\bar{B}_1$. The reason for this is that the increase in income in the depreciating country caused by the initial improvement in the trade balance leads to an increase in the depreciating country's imports, while the decline in income in the rest of the world leads to a decline in the depreciating country's exports; these adjustments cancel out a portion of the improvement in the trade balance initially produced by the depreciation.

The expansionary effects of depreciation in the country engaging in it are also likely to raise the domestic price level. This tendency may be felt in several ways:

1. The fall in imports that results directly from depreciation will cause a shift in demand toward domestically produced goods that compete with imports, and this will cause a rise in the prices of such goods.
2. The rise in demand for exports is likely to cause a rise in domestic prices of export goods. Since in most instances these goods are also used domestically, such increases affect the domestic price structure.
3. The increase in competition for factors of production will tend to push up factor prices.
4. If wages are negotiated through collective bargaining and are affected by the cost of living through escalator clauses or other arrangements, wage rates will rise. This is especially likely in the case of a country which is heavily dependent on imports for consumption.

The internal effects of depreciation on employment and the price level can be resisted by the adoption of suitably restrictive domestic fiscal and monetary policies. Indeed, if such policies are not adopted, the beneficial effect of depreciation on the balance of payments may be undermined by the rise in domestic income and prices.

Exchange depreciation may be used, and indeed has on occasion been used, primarily as a means of stimulating domestic economic activity. Perhaps the clearest instance of this was in 1933–34 when the United States depreciated the dollar by raising the price of gold from $20.67 to $35. In this case, depreciation was not necessary for balance-of-payments reasons, since the U.S. balance of payments was in surplus at the time. The objective was rather to induce recovery from the Great Depression. While the depreciation did provide some stimulus to the domestic economy, it imposed heavy deflationary pressure on the rest of the world at a time when this was especially undesirable. The use of exchange depreciation to stimulate the domestic economy in such circumstances is a prime example of a "beggar-my-neighbor" policy, discussed in Chapter 22, and is now recognized as a very undesirable action. Domestic monetary and fiscal policy, not exchange depreciation, should be used as a means of stimulating the domestic economy.

Chapter 24 PROBLEMS OF THE INTERNATIONAL MONETARY SYSTEM

In the last few years, a number of countries, most notably the United States and the United Kingdom, have experienced serious balance-of-payments problems which have confronted them with extremely difficult policy choices. Protracted U.S. and U.K. deficits have imposed severe strains on the international monetary system and have exposed certain internal contradictions in the system as presently constituted, which have led some economists and others to question its stability and viability. In this chapter we shall discuss the problems of adjustment under the present system, together with possible reforms that might be adopted to improve it.

THE WEAKNESSES OF THE SYSTEM

The period of over 20 years that has elapsed since the establishment of the International Monetary Fund has been, with only a very few minor lapses, an era of remarkable world prosperity and vigorous economic expansion. Moreover, one of the major factors underlying the prosperity of the period has undoubtedly been the steady and remarkable increase in international trade. Total world exports increased from $53.3 billion in 1948 to $212.9 billion in 1968, for an increase of 7.2 percent per year. Any international monetary system that could sustain such a period of world prosperity and expanding trade obviously has important merits. Nevertheless, it is clear that the present system also has serious weaknesses; indeed, these weaknesses have been increasingly exposed to view by the very expansion that the system has made possible.

Lack of an Adequate Adjustment Mechanism

As the analysis of Chapter 23 brings out, there are basically only three ways of eliminating a balance-of-payments deficit or surplus: (1) by corrective changes in domestic incomes, prices, and interest rates; (2) by direct regulation of trade and capital movements through the use of such devices as tariffs, import quotas, export subsidies, or exchange controls; and (3) by corrective adjustments in exchange rates. One of the most serious difficulties at the present time is that there is an extreme reluctance to use any of these methods of restoring balance-of-payments equilibrium, with the result that chronic deficits and surpluses tend to persist for long periods of time without correction.

The use of internal price and income changes as a means of maintaining or restoring balance-of-payments equilibrium is ruled out in most cases because countries insist on their right to use the instruments of monetary and fiscal policy to achieve purely domestic objectives. Indeed, this right is often viewed as the essence of national sovereignty in the field of economic policy. Thus, a country having a balance-of-payments deficit is understandably reluctant to employ a restrictive monetary and fiscal policy which would result in a higher than optimal unemployment rate; similarly, a country experiencing a surplus is very hesitant to adopt an expansionary monetary and fiscal policy which would cause prices to rise more rapidly than would be acceptable from a domestic standpoint. Indeed, in a democratic country, a government which was willing to accept a sharp increase in unemployment or a rapid inflation in order to correct a deficit or surplus would generally incur the anger of the electorate and run a grave risk of defeat at the polls.

It should not be concluded that no adjustments are ever made in monetary and fiscal policy for reasons related to the balance of payments. First, as explained earlier, balance-of-payments deficits and surpluses set in motion automatic corrective adjustments in internal demand which work toward the restoration of equilibrium unless deliberately counteracted by adjustments in monetary and fiscal policy. While countries commonly desire to neutralize these corrective forces whenever they work in a direction inconsistent with the achievement of domestic employment and price level objectives, the available instruments of fiscal and monetary policy are often cumbersome to employ and less than perfectly efficient, so that it is frequently not possible to neutralize the automatic corrective forces entirely. Second, even though a country may want to tailor its policies entirely for the achievement of domestic objectives, its international reserve position may be such as to force it to pay some attention to the balance of payments.

There is an important lack of symmetry in this regard between the position of a deficit country and that of a surplus country. The size of a country's reserves, together with its ability to borrow from the IMF or other sources, sets a fairly definite upper limit on the cumulative magnitude of the deficit it will be able to finance. Thus, if it experiences a sequence of deficits that depletes its reserves substantially, it will come under increasing pressure to attempt to correct its deficit by adopting a less expansionary monetary and fiscal policy than it would like to employ on the basis of domestic considerations alone. On the other hand, a country experiencing a sequence of surpluses that are adding to its reserves will not feel an equivalent pressure to adopt a more expansionary policy than would be dictated by domestic conditions. The reason for this asymmetry is that a country can only run its reserves down to zero through a series of deficits, whereas there is no definite upper limit on its ability to accumulate reserves through a sequence of surpluses. For this reason, the present international monetary system contains an inherent deflationary bias: the pressure on deficit countries to adopt deflationary policies is greater than the pressure on surplus countries to adopt inflationary measures. Conservatives, who often attach a heavy weight to price stability as an objective, commonly feel that the "discipline" that the balance of payments exerts on deficit countries to adopt restrictive policies is a valuable safeguard against excessive inflation. On the other hand, those who, like the present writer, attach a heavy weight to high employment as a goal and feel that in a world of Phillips curves some secular rise in the price level is unavoidable if high employment is to be sustained, view the deflationary bias exerted by the present international monetary system as distinctly undesirable.

In general, the industrial nations are committed to the principle of free movement of goods and capital as a means of achieving maximum efficiency in the use of world resources, and considerable progress has been made during the postwar era in moving toward this objective through the lowering of tariffs and the dismantling of exchange controls. As a matter of principle, this commitment rules out direct interference with trade and capital movements through the use of such devices as tariffs, quotas, exchange controls, and capital controls as means of dealing with balance-of-payments deficits. Nevertheless, there have been occasions in the last few years in which countries—including Canada, the United Kingdom, and the United States—have felt compelled to resort to such measures as a means of dealing with deficits on the ground that they were less objectionable than the other alternatives available. Thus, the United States has employed such restrictions as the "tying" of foreign aid (i.e., requiring that aid recipients spend the aid funds in the United States even though their needs can be met at a lower cost by buying elsewhere); discrimination in favor of U.S. suppliers in the placing of military contracts; and limitations, first voluntary and later mandatory, on direct investment overseas by U.S. corporations. All of these measures are inconsistent with the avowed U.S. objective of free international trade and capital movements.

Although, as explained earlier, the IMF makes specific provision for exchange parity adjustments as a means of dealing with a "fundamental disequilibrium" in the balance of payments, the advanced industrial countries have demonstrated an extreme reluctance to resort to devaluation as a means of correcting deficits except as a last resort. There are several reasons for this. In the first place, frequent adjustments of exchange parities among the leading currencies are almost certain to undermine the stability of the whole system by encouraging speculation against a country's currency whenever the country experiences a deficit. If the pound is at its lower support point of \$2.376 (as it would probably be if Britain were experiencing a large deficit) and if the parity were then reduced to \$2, an investor who was foresighted enough to shift into dollars before the devaluation would realize a profit of 18.8 percent on his investment.[1] On the other hand, in the event devaluation did not take place, he would stand to lose very little, since he could always shift back into pounds at an exchange rate which would be unlikely to rise appreciably above \$2.376 as long as Britain continued to experience a deficit, and which could not in any case rise higher than \$2.424.

In other words, the present system gives speculators something very close to a one-way option and creates an incentive to speculate on even a mild possibility that a currency will be devalued. And, of course, speculation against a currency generates a capital outflow that adds to the deficit, thereby increasing the drain on the country's reserves and making more probable the devaluation the speculators are hoping for. The more frequently devaluation is actually resorted to, the more credible it becomes in deficit situations and the more likely it is that the self-generating process described above will get under way. Thus, in considering a devaluation, a nation which is a responsible member of the international community must pay considerable attention to the possible ramifications of its action in increasing the instability of the whole international monetary system. Moreover, there is a possibility that devaluation by one major country may directly lead to devaluation by other countries who are desirous of maintaining their competitive positions in international trade. If a country devalues and its competitors follow suit by devaluing in approximately equal degree, the result is likely to be a substantial loss of confidence in the stability of the system with little or no balance-of-payments gain to the country initiating the sequence of devaluations.

The reader will recall from our earlier discussion that if the full potential benefits of the devaluation to the balance of payments are to be achieved, restrictive fiscal and monetary measures will need to be taken to prevent the

---

[1] To illustrate, suppose a British investor sold £10,000 for \$23,760 (10,000 × 2.376) before the devaluation. After the devaluation of the pound to \$2, he would be able to buy £11,880 with his \$23,760 (23,760 ÷ 2 = 11,880). Thus, he would have earned a profit of £1,880 on an investment of £10,000, for a return of 18.8 percent. What annual rate of profit this would represent would depend on the time period involved. If the investor had been prescient enough to sell sterling only a week before it was devalued, the annual rate of return would be 977.6 percent (18.8 percent × 52 weeks). There might be some loss of interest income if interest rates in New York were lower than in London, because the investor would have to forego the interest on £10,000 in London during the period, earning instead the interest on \$23,760 in New York. But this loss of interest would be trivial in relation to the profit on the turn of the exchange. A foreign (say American) investor could speculate against the pound, for example, by borrowing £10,000 in London, selling the pounds for \$23,760 before devaluation, and buying the necessary £10,000 to repay the loan for \$20,000 after devaluation. This would leave him with a profit of \$3,760. Again, there would be a small offset for net interest cost if he had to pay a higher interest rate on the loan in London than he could earn by investing the \$23,760 in New York.

improvement in the balance of payments from increasing aggregate internal demand. In addition, some means of restraining the zeal of trade unions in collective bargaining may be needed to minimize the upward pressure on wages resulting from the effect of devaluation on living costs. If such harsh and generally unpopular supporting domestic policies are not adopted, there is a serious danger that the potential benefits of devaluation will be frittered away through rising prices and costs, with a minimal improvement in the balance of payments. Furthermore, if devaluation merely leads to domestic inflation without significant improvement in the balance of payments, the country may find itself in a more precarious position than existed before the devaluation was undertaken, since an unsuccessful devaluation is likely to be taken by speculators as an indication that another dose of the same medicine will be applied. The result is likely to be a large speculative outflow of capital which may itself force a repeat performance of the devaluation.

Considering that devaluation is a threat to the stability of the entire international monetary system, that there is always danger of retaliation by other countries, that a successful devaluation requires the application of domestic policies that are likely to be both unpleasant and unpopular, and that an unsuccessful devaluation may leave a country in an even more precarious position than it started from, it is not surprising that devaluation is not a popular measure for dealing with a balance of payments deficit and is not likely to be resorted to until almost all other alternatives have been exhausted.

Instability of the System

As indicated above, the present arrangements with respect to exchange rates have created a situation in which there is very little risk involved in speculating against the currency of a country whose balance of payments is in deficit. The resulting speculative dangers not only militate against the use of exchange rate adjustments, as pointed out above, but also make the system generally vulnerable to speculative attack and potentially unstable. The lack of any generally acceptable way of restoring equilibrium means that when a structural change throws a country's balance of payments into deficit, the deterioration of the country's reserve position may continue for a considerable period. The decline in reserves—or, in the case of a reserve currency country such as the United States, the increase in its liquid liabilities—may cause a progressive decline in confidence in its exchange parity, with the result that the danger of a speculative attack on its currency increases. Not only foreign holders of claims denominated in the currency but domestic holders as well may sell the currency for other currencies in order to profit from an expected devaluation. Thus, the resources involved in a speculative attack could reach astronomical proportions, thereby necessarily exhausting the country's reserves and forcing a crisis devaluation, which might have a devastating effect on the whole international monetary system.

Inadequate Growth of Reserves

As indicated earlier, the dollar value of world exports increased at an average rate of 7.2 percent per year from 1948 to 1968. During the same period, total international monetary reserves increased from $48 billion to $76.6 billion, for an average increase of only 2.4 percent per year. Thus, reserves grew only one third as rapidly as trade. It is not easy to say how large reserves should be at any time or how rapidly they should increase. Official reserves are not, of course, used to finance private trade. Their function is rather to enable countries to finance balance-of-payments deficits. Reserves will need to be larger the greater and more frequent are the shocks to the system that generate deficits and the less effective and quick-acting is the adjustment mechanism for eliminating deficits. It seems reasonable to suppose, on this basis, that to provide a given degree of protection reserves should grow about as fast as trade. If reserves grow substantially less rapidly than trade, as has been the case during the postwar period, deficit countries are likely to come under increasing pressure to adopt restrictive domestic monetary and fiscal policies or to impose trade restrictions in order to correct deficits. Thus, inadequate growth of reserves can impose either deflation or increasingly illiberal trade policies on the world economy.

For these reasons, the fact that reserves have been growing much less rapidly than trade is a cause for concern. In addition, the composition of reserves has been changing in a direction that seems undesirable. Gold production has increased very slowly during the postwar period, having been held back by the constant $35 price of gold in the face of rising costs of production. At the same time, the demand for gold in industrial uses and, in some parts of the world, as a form in which to hold wealth has risen. The result is that the flow of gold into monetary reserves has declined: the stock of monetary gold grew at only 0.8 percent per year from 1948 to 1968. The foreign exchange component of reserves—especially in the form of dollar holdings—has grown at a rate of 4.3 percent per year. The resulting increase in U.S. liquid liabilities to foreign central banks, combined with the decline in the U.S. gold stock, has weakened confidence in the continued convertibility of the dollar into gold, thereby contributing to the increasing instability of the international monetary system. In addition, further growth of the dollar component of reserves is dependent on a continuation of U.S. deficits. As a result of increases in IMF quotas, reserve positions in the IMF have increased quite rapidly, although they still constitute a very minor portion of total reserves.

The most troublesome aspect of the situation is that the growth of world reserves has, up to now, been primarily dependent on the vagaries of gold production and U.S. balance-of-payments deficits. There has been no systematic provision for the creation of reserves at a rate judged to be appropriate to meet the needs of a growing world economy.

Summary

It does not strain reality very much to describe the world economy we seek to achieve as one in which (1) there is complete freedom in the international movement of goods and capital, (2) exchange rates are fixed within narrow limits, and (3) individual countries are left free to use monetary and fiscal policy to attain their domestic price-level and employment goals. The system is subject to the constraint that each country possesses a limited supply of reserves with which to cover deficits in its balance of payments.

The trouble with such a system is that it is inconsistent. When a country's balance of payments is thrown into deficit by technological developments, by a change in tastes, by the requirements of a foreign policy involving heavy overseas expenditures, or simply by the divergent trends in real growth or in price levels that are inherent in independent and uncoordinated domestic policies, no corrective measures are possible that do not conflict with the basic principles of the system. It is true that legal provision is made for exchange-parity adjustments, but such adjustments subject the system to severe shocks and are therefore seldom resorted to except in crisis situations. The consequence is that there is no assurance that the deficit will be eliminated before the country's limited supply of reserves is used up. And the difficulty is further compounded by the fact that the progressive dwindling of reserves is itself likely to create increasing expectations of devaluation, possibly leading to a self-generating speculative outflow of capital which can plunge the whole system into a crisis.

WAYS OF IMPROVING THE SYSTEM

Since the weaknesses in the present system are serious and since there has been much debate concerning possible improvements, we shall attempt to sketch briefly the main approaches that might be taken to reform of the system.

Increasing the Supply of Reserves

One way to improve the system is to relax the constraint within which it operates by increasing the supply of reserves available to cover deficits when they arise. This gives the balance of payments more time to right itself before deficit countries are placed under pressure to adopt undesired measures to deflate their economies or to interfere with the freedom of trade. Moreover, as pointed out above, under the arrangements that have prevailed the supply of reserves has been governed largely by the vagaries of gold production and of U.S. balance-of-payments deficits which increase foreign holdings of dollars. As a consequence, reserves have been growing much less rapidly than trade in recent years, with the result that the constraint has been getting tighter. For some time, it has been clear to nearly everyone that some means of creating reserves in an orderly and systematic way is needed if this pernicious tendency is to be stopped.

In September 1967, the member nations of the IMF, meeting in Rio de Janiero, approved the outline of a plan which had been agreed upon earlier by representatives of Switzerland and the Group of Ten countries. This plan establishes within the IMF a new facility based on so-called special drawing rights (SDR's). SDR's are to be created in accordance with the world's estimated need for reserves and allocated to all participating IMF members in proportion to their IMF quotas. An SDR has a value equal to that of one U.S. dollar, and a member country having a balance-of-payments deficit may use its SDR's to obtain convertible currencies from other countries. These currencies can then be used to finance the country's deficit. Participating countries having balance-of-payments surpluses or especially strong reserve positions agree to accept SDR's, subject to certain limitations, from other members needing to finance deficits. Thus, SDR's constitute a genuine addition to world reserves and might even in the course of time supplant gold in this respect.[2]

Several steps were required before the plan could be put into operation. First, it was embodied in draft amendments to the IMF Articles of Agreement which were then submitted to all IMF member countries for ratification. By July 1969, the draft amendments had been ratified by the necessary 60 percent of the member countries representing 80 percent of total IMF voting power, thereby bringing the amendments into force for all 111 member countries. The next step required each member country wishing to participate to certify its acceptance of the obligations connected with participation. Not until 75 percent of all member countries had completed such certification would the plan go into effect. This step was completed in August 1969. In September 1969 the managing director and the executive directors of the IMF made a formal proposal for allocations of SDR's over the ensuing three years. This proposal was discussed at the annual meeting of the board of governors of the IMF in September and was approved by the necessary 85 percent majority of the voting power of the participants.[3] The

[2] For an especially lucid explanation of the SDR plan, see Fritz Machlup, *Remaking the International Monetary System: The Rio Agreement and Beyond* (Committee for Economic Development, 1968).

[3] The requirement that allocations must be approved by an 85 percent majority gives both the Common Market countries as a group, with 17 percent of the voting power of the IMF, and the United States, with 22 percent, veto power over allocations.

first allocation of $3.5 billion was made on January 1, 1970; this is to be followed by allocations of $3 billion each on January 1, 1971, and January 1, 1972. Allocations will be made in ensuing years when the managing director of the IMF, after consultation with member countries, concludes that there is a need to supplement existing international reserves and that there is strong support among the IMF membership for the creation of SDR's. Future allocations, like the initial ones, must be approved by an 85 percent majority of the voting power of the IMF.

While the SDR plan is potentially capable of improving the international monetary system substantially, there is little reason to think that it will change the system in a fundamental way. Even with such a plan in operation the supply of reserves available to a deficit country will be limited, and, in the absence of corrective measures to end the deficit, the same speculative difficulties that have characterized the system in the past can be expected to make their appearance as a country's supply of reserves becomes depleted. While the establishment of an orderly procedure for creating additional reserves is a major reform, it is difficult to see how the inherent deficiencies of the present system can be rooted out without providing some adjustment procedures that will serve to correct underlying trade imbalances.

International Coordination of Economic Policies

Another way of strengthening the international monetary system would be to establish some "rules of the game" regarding the responsibilities for deficit and surplus countries to make appropriate adjustments in their domestic monetary and fiscal policies. That is, deficit countries might agree to adopt somewhat more restrictive domestic policies than they would otherwise prefer as their contribution toward the restoration of international balance, while surplus countries would correspondingly agree to adopt somewhat more expansionary policies than they would otherwise prefer. That is, deficit countries would accept greater than optimal unemployment and surplus countries greater than optimal inflation.

Such an agreement on the "rules of the game" would certainly be exceedingly difficult to negotiate. There would be knotty problems relating to the optimal division of the burden of adjustment between deficit and surplus countries, which would be rendered more difficult by the fact that the magnitude of the total problem would be difficult to gauge and would vary from case to case, depending on the nature and magnitude of the initial disequilibrating force. Moreover, given the constantly shifting forces affecting the economies of different countries and the relations between them, together with the less than perfect understanding of the effects of various instruments of economic policy, it would be difficult to tell whether the participating countries were abiding by the rules. Finally, it would be almost impossible for one political administration in a country to commit a future and different political administration to abide by the rules.

It may be noted that the limiting case of policy coordination is full currency unification by the participating countries.[4] As a necessary condition for currency unification to be workable, the participating countries would have to give up their sovereign authority to conduct independent monetary and fiscal policies directed at internal price and employment goals. Such policies would have to be conducted by a centralized monetary and fiscal authority charged with responsibility for internal stability for the group of countries as a whole. This centralized authority, in order to carry out its responsibilities effectively, would need to have sole power to regulate the supply of money and credit and to levy certain taxes and control certain categories of government expenditures. Under such an arrangement, internal price and income changes and interest rate adjustments would take care of the balance-of-payments problems of the individual countries. The arrangement would be very similar to the internal monetary system of the United States, and balance-of-payments problems would presumably no longer be a matter of concern.

The price that would have to be paid for this arrangement would be the loss of sovereignty over economic policy by the individual participating countries; in this respect, their position would become similar to that of individual states or Federal Reserve districts in the United States. Thus, individual countries would no longer be able to choose their optimal levels of internal demand but would have to accept the levels that were associated with the overall policies judged by the central authorities to be appropriate for the group as a whole. Some countries might find themselves in the position of depressed areas—a position very similar to that of a state like West Virginia in the U.S. federal system. The central government could, of course, alleviate localized distress by programs of expenditures and tax incentives aimed at the stimulation of production and employment in depressed areas, just as the federal government is able to do in the United States.

Since the power to create money is perhaps the fundamental element of national sovereignty, it would be difficult for the participating countries to carry out independent military and foreign policies, especially those involving heavy expenditures overseas for national security and economic aid to underdeveloped countries. With monetary sovereignty eliminated, the other elements of national sovereignty would probably wither away, with the corresponding powers being shifted to the central government. Thus, while currency unification might prove a satisfactory solution to the problems of

[4] For a fuller discussion of this and related matters, see Warren L. Smith, "Are There Enough Policy Tools?" *American Economic Review*, Vol. LV, May 1965, pp. 208–20, reprinted in W. L. Smith and R. L. Teigen (eds.), *Readings in Money, National Income, and Stabilization Policy* (2d ed.; Homewood, Ill.: Richard D. Irwin, Inc., 1970).

economic policy, it would require, directly and indirectly, such a sweeping surrender of the accepted and widely revered prerogatives of national sovereignty that it is hard to believe that it would be acceptable at the present time to many countries.

Altering the Mix of Monetary and Fiscal Policies[5]

One possibility that has received considerable attention is the sophisticated use of the "mix" of fiscal and monetary policies by individual countries to achieve internal and external stability simultaneously. The general idea as applied to the United States can be explained quite simply in the form of three propositions (which as we shall see later are subject to important qualifications):

1. If we take levels of income and prices in the rest of the world as determined by the policies followed by other countries and unaffected by developments in the United States, the U.S. balance of trade in goods and services will be determined by the level of U.S. aggregate demand. The reason is that aggregate demand will determine U.S. income and price levels, which, with foreign incomes and prices taken as given, will in turn determine U.S. imports and exports.

2. Any desired level of aggregate demand in the United States can be achieved by various combinations of fiscal and monetary policy. We can, for example, have the same level of aggregate demand with either (*a*) an easy fiscal policy involving a substantial budget deficit combined with a tight monetary policy and high interest rates, or (*b*) a tight fiscal policy involving a budget surplus combined with an easy monetary policy and low interest rates.

3. High interest rates in the United States (relative to those abroad, which are taken as given) will attract an increased net inflow of capital into the United States. Thus, if at the level of aggregate demand that is needed to achieve the employment/price-behavior target—e.g., point *N* on the modified Phillips curve in Figure 17–1[6]—the balance of trade in goods and services shows a deficit, this level of aggregate demand should be achieved with a policy mix involving sufficiently high interest rates to attract an inflow of capital large enough to cover the deficit on trade account. Conversely, if the balance of trade shows a surplus at the target level of aggregate demand, this level of demand should be achieved with a policy mix involving sufficiently low interest rates to generate a capital outflow equal to the trade surplus.

---

[5] For a more extensive treatment of some of the matters discussed here, see R. M. Stern, *The Balance of Payments: Theory and Economic Policy* (Chicago: Aldine Press, in press), chap. 9.

[6] See page 38.

It will be useful to work out these ideas somewhat more fully by developing a static linear model of the relationships involved. The model we shall use includes the following nine equations:

$$Y = C + I + G^* + \bar{X} - P \quad (1)$$
$$C = cY_d + C_0 \quad (2)$$
$$I = iY_d - vr + I_0 \quad (3)$$
$$Y_d = Y - T \quad (4)$$
$$T = xY + T_0 \quad (5)$$
$$M^* = kY - mr + M_0 \quad (6)$$
$$P = pY + P_0 \quad (7)$$
$$F = br + F_0 \quad (8)$$
$$B = \bar{X} - P + F \quad (9)$$

Here Y is GNP; C, personal consumption expenditures; G^*, government purchases of goods and services; $\bar{X}$, exports of goods and services; P, imports of goods and services; Y_d, disposable (aftertax) income; T, net taxes (i.e., taxes minus transfer payments); M^*, the stock of money; r, the interest rate; F, the net inflow of capital into the country; and B, the surplus or deficit in the balance of payments on an official reserve transactions basis. Equation 1 is the definition of GNP from the expenditure side for an open economy, as explained in Chapter 2. Equations 2, 3, and 5 are the consumption function, investment function, and tax function, respectively. Equation 4 defines disposable income. Equation 6 defines the equilibrium of the supply of money with the demand for money. Equation 7 specifies imports of goods and services as a function of GNP. Equation 8 indicates that the net inflow of capital is a function of the interest rate. Equation 9 defines the balance-of-payments surplus or deficit as the algebraic sum of the trade balance $(\bar{X} - P)$ and the capital flow balance (F).

The model makes the simplifying asumption that prices remain constant, i.e., that changes in aggregate demand cause changes in output and employment with no effect on prices. The symbols C_0, I_0, T_0, M_0, P_0, and F_0 are the constant terms in the consumption function, investment function, tax function, money-demand function, import function, and capital-inflow function, respectively. Exports $(\bar{X})$ are assumed to be constant, determined by the (given) income and price level in the rest of the world. M^* and G^* are taken to be policy parameters that are changed to implement monetary policy and fiscal policy, respectively. (Fiscal policy could alternatively be implemented by changing taxes, but for simplicity, this possibility is ruled out.) Thus, for given values of the policy parameters, M^* and G^*, the nine equations of the model are sufficient to determine the nine endogenous variables, Y, C, I, P, Y_d, r, T, F, and B.

It is possible to simplify this system considerably. First, let us substitute Equations 2, 3, 4, 5, and 7 into Equation 1. Simplifying and solving for r, we obtain the *equation of the IS curve*

$$r = -\frac{1}{v}[1 - (c + i)(1 - x) + p]Y + \frac{1}{v}[C_0 + I_0 - (c + i)T_0 - P_0 + \bar{X} + G^*] \quad (10)$$

The slope of the *IS* curve is

$$\frac{\Delta r}{\Delta Y} = -\frac{1 - (c + i)(1 - x) + p}{v}$$

We shall assume that $(c + i)(1 - x) - p$, which is the marginal propensity to spend on domestically produced output, is less than unity.[7] If this is the case, it is clear that the slope of the *IS* curve is negative.

Solving equation 6 for r, we have the *equation of the LM curve*

$$r = \frac{k}{m}Y + \frac{1}{m}(M_0 - M^*) \quad (11)$$

The slope of the *LM* curve is

$$\frac{\Delta r}{\Delta Y} = \frac{k}{m}$$

which is clearly positive.

Next we substitute Equations 7 and 8 into Equation 9, obtaining

$$B = br - pY + \bar{X} - P_0 + F_0 \quad (12)$$

Setting $B = 0$ and solving for r, we obtain the equation of what may be termed the *$B = 0$ curve*

$$r = \frac{p}{b}Y + \frac{1}{b}(P_0 - \bar{X} - F_0) \quad (13)$$

This curve encompasses all of the combinations of income and the interest rate which will produce equilibrium in the balance of payments (i.e., neither a surplus nor a deficit on an official settlements basis and neither a gain nor a loss of monetary reserves).

The slope of the $B = 0$ curve is

$$\frac{\Delta r}{\Delta Y} = \frac{p}{b}$$

which is positive. The reason for the positive slope is easy to see. If the balance of payments is in balance to begin with and if income rises, imports will be increased, thereby throwing the balance of payments into deficit if

---

[7] $(c + i)(1 - x)$ is the total marginal propensity to spend out of GNP on output, including imports. If the marginal propensity to import, p, is subtracted from this, what is left is the marginal propensity to spend on domestically produced output.

the interest rate is unchanged. In order to restore balance, it will be necessary to raise the interest rate in order to attract a sufficient additional inflow of capital to offset the increased imports. Thus, the higher is the level of income the higher will the interest rate have to be to keep payments in balance.

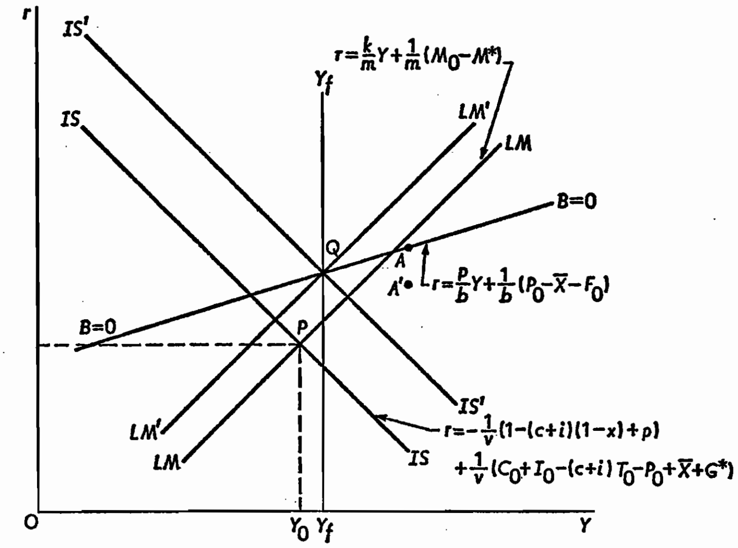
The equilibrium level of income associated with any given values of G^* and M^* can be determined by eliminating r between Equations 10 and 11 and solving the resulting equation for Y. This yields

$$Y = \frac{1}{1 - (c + i)(1 - x) + p + \dfrac{vk}{m}} \left(C_0 + I_0 - (c + i)T_0 - P_0 - \frac{v}{m} M_0 + \bar{X} + G^* + \frac{v}{m} M^* \right) \quad (14)$$

This equilibrium is depicted in Figure 24–1 by the intersection of the IS and LM curves at point P.[8]

FIGURE 24–1

Use of Monetary and Fiscal Policy to Achieve Internal and External Balance



There is, in general, no reason why the equilibrium level of income should be associated with either balance-of-payments equilibrium or full employ-

[8] For an alternative graphical presentation that leads to similar conclusions, the student is referred to R. A. Mundell, "The Appropriate Use of Monetary and Fiscal Policy for Internal and External Stability," *International Monetary Fund Staff Papers,* Vol. IX, March 1962, pp. 70–79, reprinted in Smith and Teigen, *op. cit.*

ment. In Figure 24–1, full employment income is depicted by the vertical line Y_fY_f. Thus, the equilibrium income, Y_0, is at a level below full employment. Moreover, the equilibrium point P lies below the $B = 0$ line, indicating that the balance of payments is in deficit.[9]

In order to achieve both full employment and balance-of-payments equilibrium, it is necessary to move the equilibrium position of the economy to point Q, where the $B = 0$ line intersects the full employment income line Y_fY_f. This will require that the IS curve be shifted to the right to the position IS' and the LM curve to the left to the position LM'. It is apparent that this will require a suitable increase in government purchases, G^*, and a suitable contraction of the stock of money, M^*. That is, an expansionary fiscal policy combined with a restrictive monetary policy will be required. It should be noted the position of the $B = 0$ line is not affected by either monetary or fiscal policy.

A more detailed analysis of the problem can be developed as follows. The incremental change in Y produced by incremental changes in G^* and M^* can easily be derived from Equation 14, assuming $\Delta C_0 = \Delta I_0 = \Delta T_0 = \Delta P_0 = \Delta M_0 = \Delta \bar{X} = 0$. The result is as follows:

$$\Delta Y = \frac{1}{1 - (x + i)(1 - x) + p + \frac{vk}{m}} \Delta G^* + \frac{\frac{v}{m}}{1 - (c + i)(1 - x) + p + \frac{vk}{m}} \Delta M^* \tag{15}$$

Substituting Equation 11 into Equation 12, we obtain

$$B = \left(\frac{bk}{m} - p\right) Y - \frac{b}{m} M + \bar{X} - P_0 + F_0 + \frac{b}{m} M_0$$

The incremental change in B caused by changes in Y and M^* (with $\Delta \bar{X} = \Delta P_0 = \Delta F_0 = \Delta M_0 = 0$) is

$$\Delta B = \left(\frac{bk}{m} - p\right) \Delta Y - \frac{b}{m} \Delta M^* \tag{16}$$

[9] The fact that the balance of payments is in deficit at any point below the $B = 0$ line is easily made clear by comparing points A and A' in Figure 24–1. At point A, the balance of payments is in equilibrium since the point lies on the $B = 0$ line. At point A', income is the same as at point A, so the balance of trade in goods and services is the same as it is at point A. However, the interest rate is lower and the net capital inflow is therefore smaller at A' than at A. Consequently, the overall balance of payment must be in deficit at point A'. By analogous reasoning it is clear that at any point above the $B = 0$ line the balance of payments is in surplus.

Substituting the value of ΔY from Equation 15 into Equation 16 and simplifying, we obtain the incremental change in B produced by incremental changes in G^* and M^*. The result is

$$\Delta B = \frac{\frac{bk}{m} - p}{1 - (c+i)(1-x) + p + \frac{vk}{m}} \Delta G^* - \frac{\frac{b}{m}(1 - (c+i)(1-x)) + \frac{v+b}{m} p}{1 - (c+i)(1-x) + p + \frac{vk}{m}} \Delta M^* \quad (17)$$

It will be useful at this point to bring together Equations 15 and 17.

$$\Delta Y = \frac{1}{1 - (c+i)(1-x) + p + \frac{vk}{m}} \Delta G^* + \frac{\frac{v}{m}}{1 - (c+i)(1-x) + p + \frac{vk}{m}} \Delta M^* \quad (15)$$

$$\Delta B = \frac{\frac{bk}{m} - p}{1 - (c+i)(1-x) + p + \frac{vk}{m}} \Delta G^* - \frac{\frac{b}{m}(1 - (c+i)(1-x)) + \frac{v+b}{m} p}{1 - (c+i)(1-x) + p + \frac{vk}{m}} \Delta M^* \quad (17)$$

We can select target values for ΔY and ΔB—for example, a value of ΔY that will move the economy to full employment and a value of ΔB that will move the balance of payments to equilibrium (i.e., make $B = 0$). Having selected target values of ΔY and ΔB, we can solve Equations 15 and 17 simultaneously to find the changes in the policy instruments, ΔG^* and ΔM^*, that will be needed to achieve these targets.

The easiest way to make these ideas clear is by means of a numerical illustration. Let us assume the following values: $\bar{X} = 45$, $C_0 = 0$, $I_0 = 63$, $T_0 = 40$, $M_0 = 0$, $P_0 = -50$, $F_0 = -38$, $G^* = 230$, $M^* = 190$, $c = 0.9$, $i = 0.1$, $v = 12$, $m = 15$, $k = 0.25$, $x = 0.2$, $p = 0.1$, $b = 9$. Substituting these values into Equation 14, we find that the equilibrium value of income

TABLE 24–1

Numerical Illustration of Use of Monetary and Fiscal Policies to Achieve Internal and External Balance

Variable	*How Calculated*	*Initial Equilibrium*	*New Equilibrium*	*Change*
Y	$C + I + G^* + \bar{X} - P$	1,000	1,020.0	+20.0
C	$0.9Y_d$	684	698.4	+14.4
I	$0.1Y_d - 12r + 63$	91	80.6	−10.4
G^*	Policy instrument	230	248.0	+18.0
$\bar{X}$	Given	45	45.0	...
P	$0.1Y - 50$	50	52.0	+2.0
Y_d	$Y - T$	760	776.0	+16.0
T	$0.2Y + 40$	240	244.0	+4.0
r	$0.0167Y - 0.0667M^{*a}$	4%	5%	+1%
M^*	Policy instrument	190	180.0	−10.0
F	$9r - 38$	−2	7.0	+9.0
B	$\bar{X} - P + F$	−7	0	+7.0

[a] Obtained by solving the equation, $M^* = 0.25Y - 15r$, for r.

is 1,000. The position of the economy is shown in the "initial equilibrium" column of Table 24–1.

Let us suppose that the target level of income desired for full employment is 1,020. It is apparent from Table 24–1 that the balance of payments shows a deficit of 7. The purpose of the authorities, then, is to select a set of policies that will raise income by 20 and reduce the balance-of-payments deficit by 7. Substituting the values of the parameters into Equations 15 and 17, we obtain the following:

$$\Delta Y = 2\,\Delta G^* + 1.6\,\Delta M^*$$
$$\Delta B = 0.1\,\Delta G^* - 0.52\,\Delta M^*$$

Substituting the target values $\Delta Y = 20$ and $\Delta B = 7$, we have the following pair of equations:

$$2\,\Delta G^* + 1.6\,\Delta M^* = 20$$
$$0.1\,\Delta G^* - 0.52\,\Delta M^* = 7$$

Solving these two equations simultaneously for ΔG^* and ΔM^*, we obtain $\Delta G^* = 18$ and $\Delta M^* = -10$. That is, in order to achieve full employment and balance-of-payments equilibrium, it is necessary to increase government purchases (G^*) by 18, to 248, and reduce the money stock by 10, to 180. The equilibrium position of the economy after this has been done is shown in the "new equilibrium" column of Table 24–1. We leave it as an exercise for the student to work out a diagram similar to Figure 24–1 for this numerical illustration.

We have shown by means of a relatively simple model how it may be possible, in principle, to use monetary and fiscal policy in a sophisticated way to achieve external and internal balance simultaneously. Having shown this, it is now necessary to point out a number of complications and theoretical and practical difficulties that were necessarily not evident in the above simplified presentation.

1. One minor complication is that the inflow of capital may depend on the level of income as well as the interest rate—that is, a high level of income will mean that domestic industry is profitable and this may attract capital from abroad. Allowance for this possibility would make the above model a little more complex, but it would not change the analysis in any fundamental way.

2. Instead of merely being determined by total disposable income, imports might depend on consumption and investment separately, with different marginal propensities to import for the two purposes. This would make the balance of trade depend on the composition as well as the level of aggregate demand and therefore on the mix of monetary and fiscal policies. This would not be likely to vitiate the above analysis, but if the marginal propensity to import for consumption were a great deal larger than the marginal propensity to import for investment, it is possible that the appropriate policy adjustment to raise the economy to full employment and eliminate a deficit at the same time would be to tighten fiscal policy and ease monetary policy, rather than the reverse as in the example worked out above.[10] It seems quite unlikely, however, that the difference in import propensities would be large enough to cause such a reversal of the needed policy adjustment.

3. A serious theoretical problem arises from the fact that many investors appear to hold portfolios of assets that are diversified according to the currencies in which the assets are denominated. A rise in the interest rates available in a particular country may cause these investors to shift the composition of their portfolios toward investments denominated in the currency of that country—that is, to sell assets denominated in other currencies and buy assets denominated in that country's currency. This process of changing the composition of existing portfolios may take some little time to complete, and while it is going on the balance-of-payments gains of the country that raised its interest rates may be substantial. Once it is completed, however, the gains are likely to become very much smaller, since from that point onward, the country will only benefit by getting a somewhat larger share of the current flow of financial saving which adds to asset portfolios. Thus, the balance-of-payments benefits to be obtained by raising interest

[10] A tightening of monetary policy and an easing of fiscal policy would reduce investment and increase consumption. If the marginal propensity to import for consumption were enough larger than the marginal propensity to import for investment, this might cause the trade balance to worsen by an amount greater than the increase in capital inflow caused by the rise in interest rates. If this should be the case, the appropriate adjustment would be to ease monetary policy and tighten fiscal policy.

rates may be largely transitory. This difficulty is not dealt with in the model developed above, because that model makes the unrealistic assumption that the capital *flow* per unit of time (F) depends on the interest rate. This difficulty is serious enough to cast doubt on the value of changes in the policy mix as anything more than a very temporary means of dealing with conflicts between internal and external stability. It should be understood, however, that the validity of the portfolio balance view of international capital movements on which this criticism rests has not been clearly established as correct.

4. There are grave practical difficulties in the application of the policy mix approach. In reality, we do not know with much precision the magnitude of response of either the domestic economy or the balance of payments to changes in monetary and fiscal policies. Moreover, the responses occur with lags. The balance of payments responds with lags that differ from those applicable to the domestic economy, and the lags in both responses differ as between monetary and fiscal policy. Application of the policy mix approach might on occasion involve budget deficits or surpluses that were so large or interest rates that were so high as to be unacceptable for political reasons, or because of socially undesirable impacts on particular sectors of the economy such as housing construction. Thus, lack of knowledge of the magnitude and time pattern of the economy's responses and the existence of political constraints may limit the applicability of the policy mix approach.

5. The widespread use of the policy mix approach would clearly require fairly sophisticated coordination of the monetary policies of different countries. If one country raised its domestic interest rates for balance-of-payments reasons and other countries followed suit by raising their interest rates as a means of checking domestic inflation, the world could easily get into a cycle of escalating interest rates which would have a very serious effect on worldwide capital formation and economic growth.

6. Even if the kinds of policy adjustments described above could be carried out effectively, there is some doubt about their effects on the balance of payments in the long run. If, for example, a country should have a persistent deficit in its trade accounts which it attempted to counteract by maintaining high enough interest rates to attract a compensating capital inflow, it would find its interest payments to foreign investors rising as its international indebtedness increased. Thus, in the absence of other adjustments, its trade deficit (which includes interest payments) would grow and it might have to raise interest rates further to keep its accounts in balance.

7. A policy of the type described above might have effects on productivity which would work against the elimination of the underlying balance-of-payments difficulty. For example, if a country having a trade deficit persistently maintained a tight monetary policy, this might restrain investment and capital accumulation, thereby working against the improvement in efficiency and productivity that would be needed to strengthen its competitive position in international trade. It would be possible to get around this difficulty by

introducing sufficiently strong tax incentives for investment to overcome the effects of high interest rates. Unless this was done, however, the policy might hurt the country's trade position in the long run.

We may conclude that there is some possibility of using sophisticated adjustments of the mix of monetary and fiscal policies to deal with conflicts that may arise between the goals of internal and external stability and thereby contribute something to the improvement of the balance-of-payments adjustment mechanism. But the difficulties with this approach are great, and its contribution must necessarily be quite limited.

Flexible Exchange Rates

Another way of reforming the international monetary system would be to institute a system of flexible exchange rates.[11] That is, the present IMF arrangements under which countries agree to intervene in foreign exchange markets to prevent exchange rates from fluctuating more than 1 percent on either side of an established parity would be abandoned. Exchange rates would be permitted to fluctuate on a day-to-day basis in response to changes in demand and supply in the same manner as the price of wheat adjusts in a free market. Thus, if the U.S. demand for British exports should increase from *DD* to *D′D′* as depicted in Figure 21–1, the exchange rate would simply be permitted to rise from $2 to the pound to $2.80, thereby keeping supply and demand equated in the foeign exchange market.[12]

A system of flexible exchange rates has much to recommend it. In principle at least, it would leave individual countries free to follow the fiscal and monetary policies they judge to be appropriate from the standpoint of their domestic economies. Furthermore, it would eliminate the need for international monetary reserves; since movements of the exchange rate would continuously equate the supply of and demand for foreign exchange, there would be no deficits (or surpluses) to be financed with official reserves.

Many professional economists favor flexible exchange rates. The supporters include economists of diverse socioeconomic philosophies, whose views on other aspects of economic policy often differ substantially. The general view of these economists is that we should make use of the price system and

---

[11] On the subject of flexible exchange rates, see Stern, *op. cit.*, chap. 3; Milton Friedman, "The Case for Flexible Exchange Rates," in *Essays in Positive Economics* (Chicago: University of Chicago Press, 1953), pp. 157–203; R. E. Caves, "Flexible Exchange Rates," *American Economic Review*, Vol. LIII, May 1963, pp. 120–29; J. E. Meade, "The Case for Variable Exchange Rates," *Three Banks Review*, September 1955, pp. 3–27; H. C. Wallich, Statement in *The United States Balance of Payments*, Hearings before the Joint Economic Committee, Part 3, The International Monetary System: Functioning and Possible Reform, 88th Congress, 1st Session, November 12–15, 1963 (Washington: U.S. Government Printing Office, 1963), pp. 495–99, reprinted as "In Defense of Fixed Exchange Rates," in L. S. Ritter (ed.), *Money and Economic Activity: Readings in Money and Banking* (3d ed.; Boston: Houghton-Mifflin Co., 1967), pp. 442–45.

[12] See page 445.

the free market to deal with the balance of payments just as we do in most other areas of economic life. This is a viewpoint with which the present author has a great deal of sympathy. Indeed, in the long run it is quite possible that a system of flexible exchange rates will prove to be the only practicable way out of the impasse that the present international monetary system has reached.

While flexible exchange rates are supported by many economists, there has been a decided lack of support for them up to now among central bankers, government officials, and men of affairs. Several objections have been raised. One is that the price elasticities of demand and supply of internationally traded goods may be so small that wide fluctuations in exchange rates would be necessary to maintain continuous balance-of-payments equilibrium. This should not in itself be a matter of great concern since exchange rates per se are not important. But there is some concern about the possible destabilizing effects of speculative capital movements under flexible exchange rates, and some feeling that this problem would be more serious if low elasticities of demand and supply should themselves tend to promote wild fluctuations in rates. Actually, however, it seems likely that speculation would be a much less serious problem with continuously flexible exchange rates than it is under the present system, with rates fixed at any particular time but subject to adjustments that are in fact likely to take place only when the balance of payments is far out of equilibrium and has been in that position for some time.

Another argument is that flexible rates would increase uncertainty and therefore discourage international trade and particularly investment activities which involve the commitment of funds for a considerable period of time. It is possible for traders to protect themselves against the risk of exchange rate fluctuation by use of the forward markets in foreign exchange.[13] This protection would presumably continue to be available under a system of flexible exchange rates, although its cost might be increased somewhat because the underlying risk would be greater. Thus, while international trade would probably be somewhat discouraged by flexible rates, it is doubtful whether this effect would be serious. As far as longer term investment commitments are concerned, the investor now runs the risk of being adversely affected by a currency devaluation or the application of direct controls which would prevent him from repatriating his capital and his earnings. Thus, he would probably be no worse off under fluctuating rates.

---

[13] The forward market in foreign exchange provides facilities which enable an importer who expects to have to make a payment denominated in a foreign currency to buy that currency for delivery, say, 30 days hence, at a price specified now. Or an exporter expecting to receive payment denominated in a foreign currency may sell that currency for future delivery at a specified price. The functioning of the forward market is quite complex. For a discussion, see Stern, *op. cit.*, chap. 2.

Probably much of the skepticism about a system of flexible exchange rates is simply due to the fact that we have had no experience with such a system and therefore can only deduce how it would be likely to work. There was a period in the 1930's when many countries made frequent adjustments in exchange rates and in some instances let them move in response to market forces. The fact that this was a period of worldwide depression and dwindling world trade has probably contributed to the development of an instinctive opposition to flexible rates. However, this could hardly be described as a period in which there was a system of flexible exchange rates. Indeed, there was such chaos in international finance that no international monetary system really existed. Many exchange rate adjustments in the 1930's were brought about by deliberate intervention in the exchange markets by governments to devalue their currencies as a means of increasing exports to stimulate domestic economic activity. These were glaring examples of the "beggar-my-neighbor" policies discussed in Chapter 22, and would presumably be outlawed under any system of flexible exchange rates that was adopted today.

Schemes for Limited Flexibility of Exchange Rates

Recently several proposals have been advanced for exchange rate arrangements that are intermediate between the present pegged rate system and completely flexible rates.[14] One such proposal would establish a *wider band* within which exchange rates would be permitted to move. Under this proposal, the system would work in principle as it does now, except that instead of being permitted to fluctuate only 1 percent in either direction from parity, exchange rates would be allowed to vary by, say, 5 percent either way. It is alleged that this would discourage speculation by increasing the risks borne by speculators, while still enabling traders to be reasonably certain about future rates.[15] However, the significance of this effect on speculators seems somewhat questionable. When a country experiences a severe and persistent deficit, even if it manages to avoid a devaluation, its exchange rate is likely to remain at or near the lower limit, whether this is 1 percent or

---

[14] For a useful catalog of various proposals, see J. E. Meade, "Exchange-Rate Flexibility," in *International Payments Problems* (Washington: American Enterprise Institute for Public Policy Research, 1966), pp. 67–82, reprinted in Smith and Teigen, *op. cit.*

[15] The argument that speculative risks are increased is as follows. Suppose the sterling parity is £1 = \$2.40, Britain is suffering from a severe balance-of-payments deficit with the exchange rate at its lower margin, and speculators believe sterling will be devalued. With a 1 percent band either side of parity, the exchange rate would be \$2.376. If a speculator sold sterling at this price, planning to buy it back at a lower price if a devaluation occurred, the largest loss he could suffer would be about 2 percent, which would occur if sterling was not devalued and instead rose to its upper margin of \$2.424. With a 5 percent band either side of parity, the floor would be \$2.28. If the speculator sold sterling at this price and there was no devaluation, he could conceivably suffer a loss of about 10 percent if sterling rose to its upper margin of \$2.52.

5 percent below parity. The actual risk of significant loss to speculators is likely in practice to be negligible in such circumstances, even with a somewhat wider band than now exists. It is true that with a wider band the movement of exchange rates within the band might be capable of producing some correction of the basic trade accounts by changing relative prices of exports and imports. But with divergent national trends of real growth and of price levels, it seems quite likely that many countries would soon reach a situation in which their exchange rates were at the upper or lower margin of fluctuation, at which point the system would encounter the same problems that exist under present arrangements.

A more promising proposal is for a *sliding parity*—perhaps better known as a "crawling peg." Under such a proposal, the par value in terms of the dollar of each nondollar currency would be allowed to respond to the balance-of-payments situation, but only in small increments per unit of time. For example, suppose the United Kingdom should experience a persistent deficit. The Bank of England would be allowed to lower the par value of the pound at a rate of, say, 2 or 3 percent per year in small steps taken every week or perhaps even every day. In this way, the adjustment could proceed until the exchange rate had brought the balance of payments into equilibrium. The small changes in the exchange rate would not induce speculation or disrupt trade, since in the short run—e.g., over a couple of months—the changes would be negligible.[16] If a substantial balance-of-payments disequilibrium should require a large adjustment in the exchange rate for its elimination, the adjustment might require several years to accomplish.[17] A large amount of reserves might be needed to finance the deficits during this period. The system of special drawing rights might be relied upon to provide the reserves needed to make the scheme work properly.

A major issue regarding the sliding parity proposal is whether the changes in the exchange rates should be determined automatically by the exchange market or in a discretionary way by the central bank in question.[18]

[16] It should be noted, however, that if a country had a large deficit which was expected to require a persistent devaluation of its currency over a considerable period, an appropriate adjustment of its interest rates relative to those prevailing in other countries might be needed to prevent an outflow of capital. For example, suppose the sterling exchange rate was expected to decline at a rate of 2 percent per year for some time ahead. In such a situation, interest rates in London would need to be about 2 percent higher than in New York in order to counteract the tendency for capital to move from the United Kingdom to the United States to take advantage of the anticipated appreciation of the dollar relative to sterling.

[17] For this reason, the sliding parity scheme would have a greater chance of functioning successfully if there were an initial realignment of parities designed to achieve something approximating a set of equilibrium exchange rates before the scheme was put into operation.

[18] An automatic scheme might be established along the following lines: On any given day the exchange rate might be permitted to fluctuate 1 percent either side of parity with the central bank agreeing to intervene to prevent fluctuations outside this range, as is the case under the present system. At the same time, however, a new parity

Countries anticipating balance-of-payments deficits and depreciating exchange rates tend to prefer the automatic system, while countries anticipating surpluses and appreciating rates lean toward the discretionary arrangement, since they are reluctant to see the trade advantages of an undervalued currency wiped out by market-induced rises in their exchange rates. This is not to say, of course, that there is anything like unanimous agreement concerning the desirability of *any* arrangement for increasing the flexibility of exchange rates. But it does appear that, as the difficulties of the present system have become increasingly apparent, there has been more interest, even in official circles, in exploring the possibilities of achieving increased flexibility.

FINAL COMMENT

Considerable progress has been made in the last few years in strengthening the international monetary system and improving its functioning. The biggest forward step has been the development, adoption, and the beginning of the implementation of the system for special drawing rights. With that system, the world now has the means to create additional monetary reserves as they are needed to support the growth of income and trade. Having achieved this important improvement in the international monetary system, the next step is to do something to strengthen the balance-of-payments adjustment mechanism. The most promising way to accomplish this objective is to move in the direction of greater flexibility of exchange rates. While there is still great opposition to completely flexible rates, there appears recently to have been a greater willingness to consider proposals for limited flexibility such as the sliding parity schemes. There are also some ways of strengthening the adjustment mechanism that were not discussed above, such as a code of rules governing the use of variable import duties and export subsidies as a means of correcting balance-of-payments deficits and surpluses. However, approaches of this kind are basically objectionable on efficiency grounds, are subject to possible abuses, and in any case, would be very difficult to negotiate on an acceptable basis.

would be established each day equal to the average of the actual market exchange rates that had prevailed during the preceding 365 days. Under such a scheme, with a 1 percent band either side of parity and a one-year moving average for establishing daily parities, the exchange rate could depreciate by, at most, 2 percent in a year's time if a country had a persistent deficit (or appreciate by 2 percent if a country had a persistent surplus). The amount of permitted parity movement could be increased under such a scheme either by widening the band or by shortening the period of the moving average.

Index

INDEX

This book has been set in 11 and 10 point Garamond #3, leaded 1 point. Part numbers and titles and chapter numbers are in 16 point Helvetica italic. Chapter titles are in 16 point Helvetica. The size of the type page is 27 by 45½ picas.